Solution Manual for
A Course in Game Theory

VISIT…

Solution Manual for
A Course in Game Theory
by Martin J. Osborne and Ariel Rubinstein

Martin J. Osborne
Ariel Rubinstein
with the assistance of Wulong Gu

The MIT Press
Cambridge, Massachusetts
London, England

This manual was typeset by the authors, who are greatly indebted to Donald Knuth (the creator of TEX), Leslie Lamport (the creator of LATEX), and Eberhard Mattes (the creator of emTEX) for generously putting superlative software in the public domain, and to Ed Sznyter for providing critical help with the macros we use to execute our numbering scheme.

Version 1.1, 97/4/25

Contents

Preface xi

2 Nash Equilibrium 1
Exercise 18.2 (*First price auction*) 1
Exercise 18.3 (*Second price auction*) 2
Exercise 18.5 (*War of attrition*) 2
Exercise 19.1 (*Location game*) 2
Exercise 20.2 (*Necessity of conditions in Kakutani's theorem*) 4
Exercise 20.4 (*Symmetric games*) 4
Exercise 24.1 (*Increasing payoffs in strictly competitive game*) 4
Exercise 27.2 (*BoS with imperfect information*) 5
Exercise 28.1 (*Exchange game*) 5
Exercise 28.2 (*More information may hurt*) 6

3 Mixed, Correlated, and Evolutionary Equilibrium 7
Exercise 35.1 (*Guess the average*) 7
Exercise 35.2 (*Investment race*) 7
Exercise 36.1 (*Guessing right*) 9
Exercise 36.2 (*Air strike*) 9
Exercise 36.3 (*Technical result on convex sets*) 10
Exercise 42.1 (*Examples of Harsanyi's purification*) 10
Exercise 48.1 (*Example of correlated equilibrium*) 11
Exercise 51.1 (*Existence of ESS in* 2×2 *game*) 12

4 Rationalizability and Iterated Elimination of Dominated Actions 13
Exercise 56.3 (*Example of rationalizable actions*) 13
Exercise 56.4 (*Cournot duopoly*) 13

Exercise 56.5 (*Guess the average*) 13
Exercise 57.1 (*Modified rationalizability in location game*) 14
Exercise 63.1 (*Iterated elimination in location game*) 14
Exercise 63.2 (*Dominance solvability*) 14
Exercise 64.1 (*Announcing numbers*) 15
Exercise 64.2 (*Non-weakly dominated action as best response*) 15

5 Knowledge and Equilibrium 17
Exercise 69.1 (*Example of information function*) 17
Exercise 69.2 (*Remembering numbers*) 17
Exercise 71.1 (*Information functions and knowledge functions*) 17
Exercise 71.2 (*Decisions and information*) 17
Exercise 76.1 (*Common knowledge and different beliefs*) 18
Exercise 76.2 (*Common knowledge and beliefs about lotteries*) 18
Exercise 81.1 (*Knowledge and correlated equilibrium*) 19

6 Extensive Games with Perfect Information 21
Exercise 94.2 (*Extensive games with* 2×2 *strategic forms*) 21
Exercise 98.1 (*SPE of Stackelberg game*) 21
Exercise 99.1 (*Necessity of finite horizon for one deviation property*) 21
Exercise 100.1 (*Necessity of finiteness for Kuhn's theorem*) 22
Exercise 100.2 (*SPE of games satisfying no indifference condition*) 22
Exercise 101.1 (*SPE and unreached subgames*) 23
Exercise 101.2 (*SPE and unchosen actions*) 23
Exercise 101.3 (*Armies*) 23
Exercise 102.1 (*ODP and Kuhn's theorem with chance moves*) 24
Exercise 103.1 (*Three players sharing pie*) 24
Exercise 103.2 (*Naming numbers*) 25
Exercise 103.3 (*ODP and Kuhn's theorem with simultaneous moves*) 25
Exercise 108.1 (ϵ*-equilibrium of centipede game*) 26
Exercise 114.1 (*Variant of the game* Burning money) 26
Exercise 114.2 (*Variant of the game* Burning money) 27

7 A Model of Bargaining 29
Exercise 123.1 (*One deviation property for bargaining game*) 29
Exercise 125.2 (*Constant cost of bargaining*) 29
Exercise 127.1 (*One-sided offers*) 30
Exercise 128.1 (*Finite grid of possible offers*) 30
Exercise 129.1 (*Outside options*) 32

Exercise 130.2 (*Risk of breakdown*) 33
Exercise 131.1 (*Three-player bargaining*) 33

8 Repeated Games 35
Exercise 139.1 (*Discount factors that differ*) 35
Exercise 143.1 (*Strategies and finite machines*) 35
Exercise 144.2 (*Machine that guarantees v_i*) 35
Exercise 145.1 (*Machine for Nash folk theorem*) 36
Exercise 146.1 (*Example with discounting*) 36
Exercise 148.1 (*Long- and short-lived players*) 36
Exercise 152.1 (*Game that is not full dimensional*) 36
Exercise 153.2 (*One deviation property for discounted repeated game*) 37
Exercise 157.1 (*Nash folk theorem for finitely repeated games*) 38

9 Complexity Considerations in Repeated Games 39
Exercise 169.1 (*Unequal numbers of states in machines*) 39
Exercise 173.1 (*Equilibria of the* Prisoner's Dilemma) 39
Exercise 173.2 (*Equilibria with introductory phases*) 40
Exercise 174.1 (*Case in which constituent game is extensive game*) 40

10 Implementation Theory 43
Exercise 182.1 (*DSE-implementation with strict preferences*) 43
Exercise 183.1 (*Example of non-DSE implementable rule*) 43
Exercise 185.1 (*Groves mechanisms*) 43
Exercise 191.1 (*Implementation with two individuals*) 44

11 Extensive Games with Imperfect Information 45
Exercise 203.2 (*Definition of $X_i(h)$*) 45
Exercise 208.1 (*One-player games and principles of equivalence*) 45
Exercise 216.1 (*Example of mixed and behavioral strategies*) 46
Exercise 217.1 (*Mixed and behavioral strategies and imperfect recall*) 46
Exercise 217.2 (*Splitting information sets*) 46
Exercise 217.3 (*Parlor game*) 47

12 Sequential Equilibrium 49
Exercise 226.1 (*Example of sequential equilibria*) 49
Exercise 227.1 (*One deviation property for sequential equilibrium*) 49
Exercise 229.1 (*Non-ordered information sets*) 51
Exercise 234.2 (*Sequential equilibrium and PBE*) 52

Exercise 237.1 (*Bargaining under imperfect information*) 52
Exercise 238.1 (*PBE is SE in Spence's model*) 52
Exercise 243.1 (*PBE of chain-store game*) 53
Exercise 246.2 (*Pre-trial negotiation*) 54
Exercise 252.2 (*Trembling hand perfection and coalescing of moves*) 55
Exercise 253.1 (*Example of trembling hand perfection*) 55

13 The Core 59
Exercise 259.3 (*Core of production economy*) 59
Exercise 260.2 (*Market for indivisible good*) 59
Exercise 260.4 (*Convex games*) 59
Exercise 261.1 (*Simple games*) 60
Exercise 261.2 (*Zerosum games*) 60
Exercise 261.3 (*Pollute the lake*) 60
Exercise 263.2 (*Game with empty core*) 61
Exercise 265.2 (*Syndication in a market*) 61
Exercise 267.2 (*Existence of competitive equilibrium in market*) 62
Exercise 268.1 (*Core convergence in production economy*) 62
Exercise 274.1 (*Core and equilibria of exchange economy*) 63

14 Stable Sets, the Bargaining Set, and the Shapley Value 65
Exercise 280.1 (*Stable sets of simple games*) 65
Exercise 280.2 (*Stable set of market for indivisible good*) 65
Exercise 280.3 (*Stable sets of three-player games*) 65
Exercise 280.4 (*Dummy's payoff in stable sets*) 67
Exercise 280.5 (*Generalized stable sets*) 67
Exercise 283.1 (*Core and bargaining set of market*) 67
Exercise 289.1 (*Nucleolus of production economy*) 68
Exercise 289.2 (*Nucleolus of weighted majority games*) 69
Exercise 294.2 (*Necessity of axioms for Shapley value*) 69
Exercise 295.1 (*Example of core and Shapley value*) 69
Exercise 295.2 (*Shapley value of production economy*) 70
Exercise 295.4 (*Shapley value of a model of a parliament*) 70
Exercise 295.5 (*Shapley value of convex game*) 70
Exercise 296.1 (*Coalitional bargaining*) 70

15 The Nash Bargaining Solution 73
Exercise 309.1 (*Standard Nash axiomatization*) 73
Exercise 309.2 (*Efficiency vs. individual rationality*) 73

Exercise 310.1 (*Asymmetric Nash solution*) 73
Exercise 310.2 (*Kalai–Smorodinsky solution*) 74
Exercise 312.2 (*Exact implementation of Nash solution*) 75

Preface

This manual contains solutions to the exercises in *A Course in Game Theory* by Martin J. Osborne and Ariel Rubinstein. (The sources of the problems are given in the section entitled "Notes" at the end of each chapter of the book.) We are very grateful to Wulong Gu for correcting our solutions and providing many of his own and to Ebbe Hendon for correcting our solution to Exercise 227.1. Please alert us to any errors that you detect.

Errors in the Book
Postscript and PCL files of errors in the book are kept at
`http://www.socsci.mcmaster.ca/~econ/faculty/osborne/cgt/`

Martin J. Osborne
`osborne@mcmaster.ca`
Department of Economics, McMaster University
Hamilton, Canada, L8S 4M4

Ariel Rubinstein
`rariel@ccsg.tau.ac.il`
Department of Economics, Tel Aviv University
Ramat Aviv, Israel, 69978
Department of Economics, Princeton University
Princeton, NJ 08540, USA

2 Nash Equilibrium

18.2 (*First price auction*) The set of actions of each player i is $[0, \infty)$ (the set of possible bids) and the payoff of player i is $v_i - b_i$ if his bid b_i is equal to the highest bid and no player with a lower index submits this bid, and 0 otherwise.

The set of Nash equilibria is the set of profiles b of bids with $b_1 \in [v_2, v_1]$, $b_j \leq b_1$ for all $j \neq 1$, and $b_j = b_1$ for some $j \neq 1$.

It is easy to verify that all these profiles are Nash equilibria. To see that there are no other equilibria, first we argue that there is no equilibrium in which player 1 does not obtain the object. Suppose that player $i \neq 1$ submits the highest bid b_i and $b_1 < b_i$. If $b_i > v_2$ then player i's payoff is negative, so that he can increase his payoff by bidding 0. If $b_i \leq v_2$ then player 1 can deviate to the bid b_i and win, increasing his payoff.

Now let the winning bid be b^*. We have $b^* \geq v_2$, otherwise player 2 can change his bid to some value in (v_2, b^*) and increase his payoff. Also $b^* \leq v_1$, otherwise player 1 can reduce her bid and increase her payoff. Finally, $b_j = b^*$ for some $j \neq 1$ otherwise player 1 can increase her payoff by decreasing her bid.

Comment An assumption in the exercise is that in the event of a tie for the highest bid the winner is the player with the lowest index. If in this event the object is instead allocated to each of the highest bidders with equal probability then the game has no Nash equilibrium.

If ties are broken randomly in this fashion and, in addition, we deviate from the assumptions of the exercise by assuming that there is a finite number of possible bids then if the possible bids are close enough together there is a Nash equilibrium in which player 1's bid is $b_1 \in [v_2, v_1]$ and one of the other players' bids is the largest possible bid that is less than b_1.

Note also that, in contrast to the situation in the next exercise, no player has a dominant action in the game here.

18.3 (*Second price auction*) The set of actions of each player i is $[0, \infty)$ (the set of possible bids) and the payoff of player i is $v_i - b_j$ if his bid b_i is equal to the highest bid and b_j is the highest of the other players' bids (possibly equal to b_i) and no player with a lower index submits this bid, and 0 otherwise.

For any player i the bid $b_i = v_i$ is a dominant action. To see this, let x_i be another action of player i. If $\max_{j \neq i} b_j \geq v_i$ then by bidding x_i player i either does not obtain the object or receives a nonpositive payoff, while by bidding b_i he guarantees himself a payoff of 0. If $\max_{j \neq i} b_j < v_i$ then by bidding v_i player i obtains the good at the price $\max_{j \neq i} b_j$, while by bidding x_i either he wins and pays the same price or loses.

An equilibrium in which player j obtains the good is that in which $b_1 < v_j$, $b_j > v_1$, and $b_i = 0$ for all players $i \notin \{1, j\}$.

18.5 (*War of attrition*) The set of actions of each player i is $A_i = [0, \infty)$ and his payoff function is

$$u_i(t_1, t_2) = \begin{cases} -t_i & \text{if } t_i < t_j \\ v_i/2 - t_i & \text{if } t_i = t_j \\ v_i - t_j & \text{if } t_i > t_j \end{cases}$$

where $j \in \{1, 2\} \setminus \{i\}$. Let (t_1, t_2) be a pair of actions. If $t_1 = t_2$ then by conceding slightly later than t_1 player 1 can obtain the object in its entirety instead of getting just half of it, so this is not an equilibrium. If $0 < t_1 < t_2$ then player 1 can increase her payoff to zero by deviating to $t_1 = 0$. Finally, if $0 = t_1 < t_2$ then player 1 can increase her payoff by deviating to a time slightly after t_2 unless $v_1 - t_2 \leq 0$. Similarly for $0 = t_2 < t_1$ to constitute an equilibrium we need $v_2 - t_1 \leq 0$. Hence (t_1, t_2) is a Nash equilibrium if and only if either $0 = t_1 < t_2$ and $t_2 \geq v_1$ or $0 = t_2 < t_1$ and $t_1 \geq v_2$.

Comment An interesting feature of this result is that the equilibrium outcome is independent of the players' valuations of the object.

19.1 (*Location game*)[1] There are n players, each of whose set of actions is $\{Out\} \cup [0, 1]$. (Note that the model differs from Hotelling's in that players choose whether or not to become candidates.) Each player prefers an action profile in which he obtains more votes than any other player to one in which he ties for the largest number of votes; he prefers an outcome in which he ties for

[1] *Correction to first printing of book*: The first sentence on page 19 of the book should be amended to read "There is a continuum of citizens, each of whom has a favorite position; the distribution of favorite positions is given by a density function f on $[0, 1]$ with $f(x) > 0$ for all $x \in [0, 1]$."

first place (regardless of the number of candidates with whom he ties) to one in which he stays out of the competition; and he prefers to stay out than to enter and lose.

Let F be the distribution function of the citizens' favorite positions and let $m = F^{-1}(\frac{1}{2})$ be its median (which is unique, since the density f is everywhere positive).

It is easy to check that for $n = 2$ the game has a unique Nash equilibrium, in which both players choose m.

The argument that for $n = 3$ the game has no Nash equilibrium is as follows.

- There is no equilibrium in which some player becomes a candidate and loses, since that player could instead stay out of the competition. Thus in any equilibrium all candidates must tie for first place.

- There is no equilibrium in which a single player becomes a candidate, since by choosing the same position any of the remaining players ties for first place.

- There is no equilibrium in which two players become candidates, since by the argument for $n = 2$ in any such equilibrium they must both choose the median position m, in which case the third player can enter close to that position and win outright.

- There is no equilibrium in which all three players become candidates:

  - if all three choose the same position then any one of them can choose a position slightly different and win outright rather than tying for first place;
  - if two choose the same position while the other chooses a different position then the lone candidate can move closer to the other two and win outright.
  - if all three choose different positions then (given that they tie for first place) either of the extreme candidates can move closer to his neighbor and win outright.

Comment If the density f is not everywhere positive then the set of medians may be an interval, say $[\underline{m}, \overline{m}]$. In this case the game has Nash equilibria when $n = 3$; in all equilibria exactly two players become candidates, one choosing $\underline{m}$ and the other choosing $\overline{m}$.

20.2 (*Necessity of conditions in Kakutani's theorem*)

i. X is the real line and $f(x) = x + 1$.

ii. X is the unit circle, and f is rotation by 90°.

iii. $X = [0, 1]$ and

$$f(x) = \begin{cases} \{1\} & \text{if } x < \frac{1}{2} \\ \{0, 1\} & \text{if } x = \frac{1}{2} \\ \{0\} & \text{if } x > \frac{1}{2}. \end{cases}$$

iv. $X = [0, 1]$; $f(x) = 1$ if $x < 1$ and $f(1) = 0$.

20.4 (*Symmetric games*) Define the function $F: A_1 \to A_1$ by $F(a_1) = B_2(a_1)$ (the best response of player 2 to a_1). The function F satisfies the conditions of Lemma 20.1, and hence has a fixed point, say a_1^*. The pair of actions (a_1^*, a_1^*) is a Nash equilibrium of the game since, given the symmetry, if a_1^* is a best response of player 2 to a_1^* then it is also a best response of player 1 to a_1^*.

A symmetric finite game that has no symmetric equilibrium is *Hawk–Dove* (Figure 17.2).

Comment In the next chapter of the book we introduce the notion of a mixed strategy. From the first part of the exercise it follows that a finite symmetric game has a symmetric mixed strategy equilibrium.

24.1 (*Increasing payoffs in strictly competitive game*)

a. Let u_i be player i's payoff function in the game G, let w_i be his payoff function in G', and let (x^*, y^*) be a Nash equilibrium of G'. Then, using part (b) of Proposition 22.2, we have $w_1(x^*, y^*) = \min_y \max_x w_1(x, y) \geq \min_y \max_x u_1(x, y)$, which is the value of G.

b. This follows from part (b) of Proposition 22.2 and the fact that for any function f we have $\max_{x \in X} f(x) \geq \max_{x \in Y} f(x)$ if $Y \subseteq X$.

c. In the unique equilibrium of the game

$3, 3$	$1, 1$
$1, 0$	$0, 1$

player 1 receives a payoff of 3, while in the unique equilibrium of

$3,3$	$1,1$
$4,0$	$2,1$

she receives a payoff of 2. If she is prohibited from using her second action in this second game then she obtains an equilibrium payoff of 3, however.

27.2 (*BoS with imperfect information*) The Bayesian game is as follows. There are two players, say $N = \{1,2\}$, and four states, say $\Omega = \{(B,B),(B,S),(S,B),(S,S)\}$, where the state (X,Y) is interpreted as a situation in which player 1's preferred composer is X and player 2's is Y. The set A_i of actions of each player i is $\{B,S\}$, the set of signals that player i may receive is $\{B,S\}$, and player i's signal function τ_i is defined by $\tau_i(\omega) = \omega_i$. A belief of each player i is a probability distribution p_i over Ω. Player 1's preferences are those represented by the payoff function defined as follows. If $\omega_1 = B$ then $u_1((B,B),\omega) = 2$, $u_1((S,S),\omega) = 1$, and $u_1((B,S),\omega) = u_1((S,B),\omega) = 0$; if $\omega_1 = S$ then u_1 is defined analogously. Player 2's preferences are defined similarly.

For any beliefs the game has Nash equilibria $((B,B),(B,B))$ (i.e. each type of each player chooses B) and $((S,S),(S,S))$. If one player's equilibrium action is independent of his type then the other player's is also. Thus in any other equilibrium the two types of each player choose different actions. Whether such a profile is an equilibrium depends on the beliefs. Let $q_X = p_2(X,X)/[p_2(B,X) + p_2(S,X)]$ (the probability that player 2 assigns to the event that player 1 prefers X conditional on player 2 preferring X) and let $p_X = p_1(X,X)/[p_1(X,B) + p_1(X,S)]$ (the probability that player 1 assigns to the event that player 2 prefers X conditional on player 1 preferring X). If, for example, $p_X \geq \frac{1}{3}$ and $q_X \geq \frac{1}{3}$ for $X = B$, S, then $((B,S),(B,S))$ is an equilibrium.

28.1 (*Exchange game*) In the Bayesian game there are two players, say $N = \{1,2\}$, the set of states is $\Omega = S \times S$, the set of actions of each player is $\{Exchange, Don't\ exchange\}$, the signal function of each player i is defined by $\tau_i(s_1,s_2) = s_i$, and each player's belief on Ω is that generated by two independent copies of F. Each player's preferences are represented by the payoff

function $u_i((X,Y),\omega) = \omega_j$ if $X = Y =$ *Exchange* and $u_i((X,Y),\omega) = \omega_i$ otherwise.

Let x be the smallest possible prize and let M_i be the highest type of player i that chooses *Exchange*. If $M_i > x$ then it is optimal for type x of player j to choose *Exchange*. Thus if $M_i \geq M_j$ and $M_i > x$ then it is optimal for type M_i of player i to choose *Don't exchange*, since the expected value of the prizes of the types of player j that choose *Exchange* is less than M_i. Thus in any possible Nash equilibrium $M_i = M_j = x$: the only prizes that may be exchanged are the smallest.

28.2 (*More information may hurt*) Consider the Bayesian game in which $N = \{1,2\}$, $\Omega = \{\omega_1, \omega_2\}$, the set of actions of player 1 is $\{U, D\}$, the set of actions of player 2 is $\{L, M, R\}$, player 1's signal function is defined by $\tau_1(\omega_1) = 1$ and $\tau_1(\omega_2) = 2$, player 2's signal function is defined by $\tau_2(\omega_1) = \tau_2(\omega_2) = 0$, the belief of each player is $(\frac{1}{2}, \frac{1}{2})$, and the preferences of each player are represented by the expected value of the payoff function shown in Figure 6.1 (where $0 < \epsilon < \frac{1}{2}$).

State ω_1

	L	M	R
U	$1, 2\epsilon$	$1, 0$	$1, 3\epsilon$
D	$2, 2$	$0, 0$	$0, 3$

State ω_2

	L	M	R
U	$1, 2\epsilon$	$1, 3\epsilon$	$1, 0$
D	$2, 2$	$0, 3$	$0, 0$

Figure 6.1 The payoffs in the Bayesian game for Exercise 28.2.

This game has a unique Nash equilibrium $((D, D), L)$ (that is, both types of player 1 choose D and player 2 chooses L). The expected payoffs at the equilibrium are $(2, 2)$.

In the game in which player 2, as well as player 1, is informed of the state, the unique Nash equilibrium when the state is ω_1 is (U, R); the unique Nash equilibrium when the state is ω_2 is (U, M). In both cases the payoff is $(1, 3\epsilon)$, so that player 2 is worse off than he is when he is ill-informed.

3 Mixed, Correlated, and Evolutionary Equilibrium

35.1 (*Guess the average*) Let k^* be the largest number to which any player's strategy assigns positive probability in a mixed strategy equilibrium and assume that player i's strategy does so. We now argue as follows.

- In order for player i's strategy to be optimal his payoff from the pure strategy k^* must be equal to his equilibrium payoff.

- In any equilibrium player i's expected payoff is positive, since for any strategies of the other players he has a pure strategy that for some realization of the other players' strategies is at least as close to $\frac{2}{3}$ of the average number as any other player's number.

- In any realization of the strategies in which player i chooses k^*, some other player also chooses k^*, since by the previous two points player i's payoff is positive in this case, so that no other player's number is closer to $\frac{2}{3}$ of the average number than k^*. (Note that all the other numbers cannot be less than $\frac{2}{3}$ of the average number.)

- In any realization of the strategies in which player i chooses $k^* \geq 1$, he can increase his payoff by choosing $k^* - 1$, since by making this change he becomes the outright winner rather than tying with at least one other player.

The remaining possibility is that $k^* = 1$: every player uses the pure strategy in which he announces the number 1.

35.2 (*Investment race*) The set of actions of each player i is $A_i = [0,1]$. The payoff function of player i is

$$u_i(a_1,a_2) = \begin{cases} -a_i & \text{if } a_i < a_j \\ \frac{1}{2} - a_i & \text{if } a_i = a_j \\ 1 - a_i & \text{if } a_i > a_j, \end{cases}$$

where $j \in \{1,2\} \setminus \{i\}$.

We can represent a mixed strategy of a player i in this game by a probability distribution function F_i on the interval $[0,1]$, with the interpretation that $F_i(v)$ is the probability that player i chooses an action in the interval $[0,v]$. Define the *support* of F_i to be the set of points v for which $F_i(v+\epsilon) - F_i(v-\epsilon) > 0$ for all $\epsilon > 0$, and define v to be an *atom* of F_i if $F_i(v) > \lim_{\epsilon \downarrow 0} F_i(v-\epsilon)$. Suppose that (F_1^*, F_2^*) is a mixed strategy Nash equilibrium of the game and let S_i^* be the support of F_i^* for $i = 1, 2$.

Step . $S_1^* = S_2^*$.

Proof. If not then there is an open interval, say (v,w), to which F_i^* assigns positive probability while F_j^* assigns zero probability (for some i, j). But then i can increase his payoff by transferring probability to smaller values within the interval (since this does not affect the probability that he wins or loses, but increases his payoff in both cases).

Step . If v is an atom of F_i^* then it is not an atom of F_j^* and for some $\epsilon > 0$ the set S_j^* contains no point in $(v-\epsilon, v)$.

Proof. If v is an atom of F_i^* then for some $\epsilon > 0$, no action in $(v-\epsilon, v]$ is optimal for player j since by moving any probability mass in F_i^* that is in this interval to either $v + \delta$ for some small $\delta > 0$ (if $v < 1$) or 0 (if $v = 1$), player j increases his payoff.

Step . If $v > 0$ then v is not an atom of F_i^* for $i = 1, 2$.

Proof. If $v > 0$ is an atom of F_i^* then, using Step 2, player i can increase his payoff by transferring the probability attached to the atom to a smaller point in the interval $(v-\epsilon, v)$.

Step . $S_i^* = [0, M]$ for some $M > 0$ for $i = 1, 2$.

Proof. Suppose that $v \notin S_i^*$ and let $w^* = \inf\{w : w \in S_i^* \text{ and } w \geq v\} > v$. By Step 1 we have $w^* \in S_j^*$, and hence, given that w^* is not an atom of F_i^* by Step 3, we require j's payoff at w^* to be no less than his payoff at v. Hence $w^* = v$. By Step 2 at most one distribution has an atom at 0, so $M > 0$.

Step . $S_i^* = [0,1]$ and $F_i^*(v) = v$ for $v \in [0,1]$ and $i = 1, 2$.

Proof. By Steps 2 and 3 each equilibrium distribution is atomless, except possibly at 0, where at most one distribution, say F_i^*, has an atom. The payoff of j at $v > 0$ is $F_i^*(v) - v$, where $i \neq j$. Thus the constancy of i's payoff on $[0, M]$ and $F_j^*(0) = 0$ requires that $F_j^*(v) = v$, which implies that $M = 1$. The constancy of j's payoff then implies that $F_i^*(v) = v$.

We conclude that the game has a unique mixed strategy equilibrium, in which each player's probability distribution is uniform on $[0,1]$.

36.1 (*Guessing right*) In the game each player has K actions; $u_1(k,k) = 1$ for each $k \in \{1, \ldots, K\}$ and $u_1(k,\ell) = 0$ if $k \neq \ell$. The strategy pair $((1/K, \ldots, 1/K), (1/K, \ldots, 1/K))$ is the unique mixed strategy equilibrium, with an expected payoff to player 1 of $1/K$. To see this, let (p^*, q^*) be a mixed strategy equilibrium. If $p_k^* > 0$ then the optimality of the action k for player 1 implies that q_k^* is maximal among all the q_ℓ^*, so that in particular $q_k^* > 0$, which implies that p_k^* is minimal among all the p_ℓ^*, so that $p_k^* \leq 1/K$. Hence $p_k^* = 1/K$ for all k; similarly $q_k = 1/K$ for all k.

36.2 (*Air strike*) The payoffs of player 1 are given by the matrix

$$\begin{pmatrix} 0 & v_1 & v_1 \\ v_2 & 0 & v_2 \\ v_3 & v_3 & 0 \end{pmatrix}$$

Let (p^*, q^*) be a mixed strategy equilibrium.

Step 1. If $p_i^* = 0$ then $q_i^* = 0$ (otherwise q^* is not a best response to p^*); but if $q_i^* = 0$ and $i \leq 2$ then $p_{i+1} = 0$ (since player i can achieve v_i by choosing i). Thus if for $i \leq 2$ target i is not attacked then target $i+1$ is not attacked either.

Step 2. $p^* \neq (1,0,0)$: it is not the case that only target 1 is attacked.

Step 3. The remaining possibilities are that only targets 1 and 2 are attacked or all three targets are attacked.

- If only targets 1 and 2 are attacked the requirement that the players be indifferent between the strategies that they use with positive probability implies that $p^* = (v_2/(v_1+v_2), v_1/(v_1+v_2), 0)$ and $q^* = (v_1/(v_1+v_2), v_2/(v_1+v_2), 0)$. Thus the expected payoff of Army A is $v_1 v_2/(v_1+v_2)$. Hence this is an equilibrium if $v_3 \leq v_1 v_2/(v_1+v_2)$.

- If all three targets are attacked the indifference conditions imply that the probabilities of attack are in the proportions $v_2v_3 : v_1v_3 : v_1v_2$ and the probabilities of defense are in the proportions $z - 2v_2v_3 : z - 2v_3v_1 : z - 2v_1v_2$ where $z = v_1v_2 + v_2v_3 + v_3v_1$. For an equilibrium we need these three proportions to be nonnegative, which is equivalent to $z - 2v_1v_2 \geq 0$, or $v_3 \geq v_1v_2/(v_1 + v_2)$.

36.3 (*Technical result on convex sets*) NOTE: The following argument is simpler than the one suggested in the first printing of the book (which is given afterwards).

Consider the strictly competitive game in which the set of actions of player 1 is X, that of player 2 is Y, the payoff function of player 1 is $u_1(x, y) = -x \cdot y$, and the payoff function of player 2 is $u_2(x, y) = x \cdot y$. By Proposition 20.3 this game has a Nash equilibrium, say (x^*, y^*); by the definition of an equilibrium we have $x^* \cdot y \leq x^* \cdot y^* \leq x \cdot y^*$ for all $x \in X$ and $y \in Y$.

The argument suggested in the first printing of the book (which is elementary, not relying on the result that an equilibrium exists, but more difficult than the argument given in the previous paragraph) is the following.

Let $G(n)$ be the strictly competitive game in which each player has n actions and the payoff function of player 1 is given by $u_1(i, j) = x^i \cdot y^j$. Let $v(n)$ be the value of $G(n)$ and let α^n be a mixed strategy equilibrium. Then $U_1(\alpha_1, \alpha_2^n) \leq v(n) \leq U_1(\alpha_1^n, \alpha_2)$ for every mixed strategy α_1 of player 1 and every mixed strategy α_2 of player 2 (by Proposition 22.2). Let $x^{*n} = \sum_{i=1}^n \alpha_1^n(i)x^i$ and $y^{*n} = \sum_{j=1}^n \alpha_2^n(j)y^j$. Then $x^i \cdot y^{*n} \leq v(n) = x^{*n}y^{*n} \leq x^{*n} \cdot y^j$ for all i and j. Letting $n \to \infty$ through a subsequence for which x^{*n} and y^{*n} converge, say to x^* and y^*, we obtain the result.

42.1 (*Examples of Harsanyi's purification*)[1]

a. The pure equilibria are trivially approachable. Now consider the strictly mixed equilibrium. The payoffs in the Bayesian game $G(\gamma\boldsymbol{\epsilon})$ are as follows:

	a_2	b_2
a_1	$2 + \gamma\boldsymbol{\delta}_1, 1 + \gamma\boldsymbol{\delta}_2$	$\gamma\boldsymbol{\delta}_1, 0$
a_2	$0, \gamma\boldsymbol{\delta}_2$	$1, 2$

For $i = 1, 2$ let p_i be the probability that player i's type is one for which he chooses a_i in some Nash equilibrium of $G(\gamma\boldsymbol{\epsilon})$. Then it is optimal for player 1

[1] *Correction to first printing of book*: The $\boldsymbol{\epsilon}_1(x, b_2)$ near the end of line −4 should be $\boldsymbol{\epsilon}_2(x, b_2)$.

to choose a_1 if

$$(2+\gamma\delta_1)p_2 \geq (1-\gamma\delta_1)(1-p_2),$$

or $\delta_1 \geq (1-3p_2)/\gamma$. Now, the probability that δ_1 is at least $(1-3p_2)/\gamma$ is $\frac{1}{2}(1-(1-3p_2)/\gamma)$ if $-1 \leq (1-3p_2)/\gamma \leq 1$, or $\frac{1}{3}(1-\gamma) \leq p_2 \leq \frac{1}{3}(1+\gamma)$. This if p_2 lies in this range we have $p_1 = \frac{1}{2}(1-(1-3p_2)/\gamma)$. By a symmetric argument we have $p_2 = \frac{1}{2}(1-(2-3p_1)/\gamma)$ if $\frac{1}{3}(2-\gamma) \leq p_1 \leq \frac{1}{3}(2+\gamma)$. Solving for p_1 and p_2 we find that $p_1 = (2+\gamma)/(3+2\gamma)$ and $p_2 = (1+\gamma)/(3+2\gamma)$ satisfies these conditions. Since $(p_1, p_2) \to (\frac{2}{3}, \frac{1}{3})$ as $\gamma \to 0$ the mixed strategy equilibrium is approachable.

b. The payoffs in the Bayesian game $G(\gamma\boldsymbol{\epsilon})$ are as follows:

	a_2	b_2
a_1	$1+\gamma\boldsymbol{\delta}_1, 1+\gamma\boldsymbol{\delta}_2$	$\gamma\boldsymbol{\delta}_1, 0$
a_2	$0, \gamma\boldsymbol{\delta}_2$	$0, 0$

For $i = 1, 2$ let p_i be the probability that player i's type is one for which he chooses a_i in some Nash equilibrium of $G(\gamma\boldsymbol{\epsilon})$. Whenever $\delta_j > 0$, which occurs with probability $\frac{1}{2}$, the action a_j dominates b_j; thus we have $p_j \geq \frac{1}{2}$. Now, player i's payoff to a_i is $p_j(1+\gamma\delta_i)+(1-p_j)\gamma\delta_i = p_j+\gamma\delta_i$, which, given $p_j \geq \frac{1}{2}$, is positive for all values of δ_i if $\gamma < \frac{1}{2}$. Thus if $\gamma < \frac{1}{2}$ all types of player i choose a_i. Hence if $\gamma < \frac{1}{2}$ the Bayesian game $G(\gamma\boldsymbol{\epsilon})$ has a unique Nash equilibrium, in which every type of each player i uses the pure strategy a_i. Thus only the pure strategy equilibrium (a_1, a_2) of the original game is approachable.

c. In any Nash equilibrium of the Bayesian game $G(\gamma\boldsymbol{\epsilon})$ player i chooses a_i whenever $\delta_i > 0$ and b_i whenever $\delta_i < 0$; since δ_i is positive with probability $\frac{1}{2}$ and negative with probability $\frac{1}{2}$ the result follows.

48.1 (*Example of correlated equilibrium*)

a. The pure strategy equilibria are (B, L, A), (T, R, A), (B, L, C), and (T, R, C).

b. A correlated equilibrium with the outcome described is given by: $\Omega = \{x, y\}$, $\pi(x) = \pi(y) = \frac{1}{2}$; $\mathcal{P}_1 = \mathcal{P}_2 = \{\{x\}, \{y\}\}$, $\mathcal{P}_3 = \Omega$; $\sigma_1(\{x\}) = T$, $\sigma_1(\{y\}) = B$; $\sigma_2(\{x\}) = L$, $\sigma_2(\{y\}) = R$; $\sigma_3(\Omega) = B$. Note that player 3 knows that (T, L) and (B, R) will occur with equal probabilities, so that if she deviates to A or C she obtains $\frac{3}{2} < 2$.

c. If player 3 were to have the same information as players 1 and 2 then the outcome would be one of those predicted by the notion of Nash equilibrium, in all of which she obtains a payoff of zero.

51.1 (*Existence of ESS in* 2×2 *game*) Let the game be as follows:

	C	D
C	w, w	x, y
D	y, x	z, z

If $w > y$ then (C, C) is a strict equilibrium, so that C is an ESS. If $z > x$ then (D, D) is a strict equilibrium, so that D is an ESS. If $w < y$ and $x < z$ then the game has a symmetric mixed strategy equilibrium (m^*, m^*) in which m^* attaches the probability $p^* = (z - x)/(w - y + z - x)$ to C. To verify that m^* is an ESS, we need to show that $u(m, m) < u(m^*, m)$ for any mixed strategy $m \neq m^*$. Let p be the probability that m attaches to C. Then

$$\begin{aligned} u(m,m) - u(m^*,m) &= (p - p^*)[pw + (1-p)x] - (p - p^*)[py + (1-p)z] \\ &= (p - p^*)[p(w - y + z - x) + x - z] \\ &= (p - p^*)^2(w - y + z - x) \\ &< 0. \end{aligned}$$

4 Rationalizability and Iterated Elimination of Dominated Actions

56.3 (*Example of rationalizable actions*) The actions of player 1 that are rationalizable are a_1, a_2, and a_3; those of player 2 are b_1, b_2, and b_3. The actions a_2 and b_2 are rationalizable since (a_2, b_2) is a Nash equilibrium. Since a_1 is a best response to b_3, b_3 is a best response to a_3, a_3 is a best response to b_1, and b_1 is a best response to a_1 the actions a_1, a_3, b_1, and b_3 are rationalizable. The action b_4 is not rationalizable since if the probability that player 2's belief assigns to a_4 exceeds $\frac{1}{2}$ then b_3 yields a payoff higher than does b_4, while if this probability is at most $\frac{1}{2}$ then b_2 yields a payoff higher than does b_4. The action a_4 is not rationalizable since without b_4 in the support of player 1's belief, a_4 is dominated by a_2.

Comment That b_4 is not rationalizable also follows from Lemma 60.1, since b_4 is strictly dominated by the mixed strategy that assigns the probability $\frac{1}{3}$ to b_1, b_2, and b_3.

56.4 (*Cournot duopoly*) Player i's best response function is $B_i(a_j) = (1 - a_j)/2$; hence the only Nash equilibrium is $(\frac{1}{3}, \frac{1}{3})$.

Since the game is symmetric, the set of rationalizable actions is the same for both players; denote it by Z. Let $m = \inf Z$ and $M = \sup Z$. Any best response of player i to a belief of player j whose support is a subset of Z maximizes $\mathrm{E}[a_i(1-a_i-a_j)] = a_i(1-a_i-\mathrm{E}[a_j])$, and thus is equal to $B_i(\mathrm{E}[a_j]) \in [B_j(M), B_j(m)] = [(1-M)/2, (1-m)/2]$. Hence (using Definition 55.1), we need $(1-M)/2 \leq m$ and $M \leq (1-m)/2$, so that $M = m = \frac{1}{3}$: $\frac{1}{3}$ is the only rationalizable action of each player.

56.5 (*Guess the average*) Since the game is symmetric, the set of rationalizable actions is the same, say Z, for all players. Let k^* be the largest number in Z. By the argument in the solution to Exercise 35.1 the action k^* is a best

response to a belief whose support is a subset of Z only if $k^* = 1$. The result follows from Definition 55.1.

57.1 (*Modified rationalizability in location game*) The best response function of each player i is $B_i(a_j) = \{a_j\}$. Hence (a_1, a_2) is a Nash equilibrium if and only if $a_1 = a_2$ for $i = 1, 2$. Thus any $x \in [0, 1]$ is rationalizable.

Fix $i \in \{1, 2\}$ and define a pair $(a_i, d) \in A_i \times [0, 1]$ (where d is the information about the distance to a_j) to be *rationalizable* if for $j = 1, 2$ there is a subset Z_j of A_j such that $a_i \in Z_i$ and every action $a_j \in Z_j$ is a best response to a belief of player j whose support is a subset of $Z_k \cap \{a_j + d, a_j - d\}$ (where $k \neq j$).

In order for (a_i, d) to be rationalizable the action a_i must be a best response to a belief that is a subset of $\{a_i + d, a_i - d\}$. This belief must assign positive probability to both points in the set (otherwise the best response is to locate at one of the points). Thus Z_j must contain both $a_i + d$ and $a_i - d$, and hence each of these must be best responses for player j to beliefs with supports $\{a_i + 2d, a_i\}$ and $\{a_i, a_i - 2d\}$. Continuing the argument we conclude that Z_j must contain all points of the form $a_i + md$ for every integer m, which is not possible if $d > 0$ since $A_i = [0, 1]$. Hence (a_i, d) is rationalizable only if $d = 0$; it is easy to see that $(a_i, 0)$ is in fact rationalizable for any $a_i \in A_i$.

63.1 (*Iterated elimination in location game*) Only one round of elimination is needed: every action other than $\frac{1}{2}$ is weakly dominated by the action $\frac{1}{2}$. (In fact $\frac{1}{2}$ is the only action that survives iterated elimination of *strictly* dominated actions: on the first round *Out* is strictly dominated by $\frac{1}{2}$, and in every subsequent round each of the remaining most extreme actions is strictly dominated by $\frac{1}{2}$.)

63.2 (*Dominance solvability*) Consider the game in Figure 14.1. This game is dominance solvable, the only surviving outcome being (T, L). However, if B is deleted then neither of the remaining actions of player 2 is dominated, so that both (T, L) and (T, R) survive iterated elimination of dominated actions.

	L	R
T	$1, 0$	$0, 0$
B	$0, 1$	$0, 0$

Figure 14.1 The game for the solution to Exercise 63.2.

64.1 (*Announcing numbers*) At the first round every action $a_i \leq 50$ of each player i is weakly dominated by $a_i + 1$. No other action is weakly dominated, since 100 is a strict best response to 0 and every other action $a_i \geq 51$ is a best response to $a_i + 1$. At every subsequent round up to 50 one action is eliminated for each player: at the second round this action is 100, at the third round it is 99, and so on. After round 50 the single action pair $(51, 51)$ remains, with payoffs of $(50, 50)$.

64.2 (*Non-weakly dominated action as best response*) From the result in Exercise 36.3, for any ϵ there exist $p(\epsilon) \in P(\epsilon)$ and $u(\epsilon) \in U$ such that

$$p(\epsilon) \cdot u \leq p(\epsilon) \cdot u(\epsilon) \leq p \cdot u(\epsilon) \text{ for all } p \in P(\epsilon), u \in U.$$

Choose any sequence $\epsilon_n \to 0$ such that $u(\epsilon_n)$ converges to some $\overline{u}$. Since $u^* = 0 \in U$ we have $0 \leq p(\epsilon_n) \cdot u(\epsilon_n) \leq p \cdot u(\epsilon_n)$ for all n and all $p \in P(0)$ and hence $p \cdot \overline{u} \geq 0$ for all $p \in P(0)$. It follows that $\overline{u} \geq 0$ and hence $\overline{u} = u^*$, since u^* corresponds to a mixed strategy that is not weakly dominated.

Finally, $p(\epsilon_n) \cdot u \leq p(\epsilon_n) \cdot u(\epsilon_n)$ for all $u \in U$, so that u^* is in the closure of the set B of members of U for which there is a supporting hyperplane whose normal has positive components. Since U is determined by a finite set, the set B is closed. Thus there exists a strictly positive vector p^* with $p^* \cdot u^* \geq p^* \cdot u$ for all $u \in U$.

Comment This exercise is quite difficult.

5 Knowledge and Equilibrium

69.1 (*Example of information function*) No, P may not be partitional. For example, it is not if the answers to the three questions at ω_1 are (Yes, No, No) and the answers at ω_2 are (Yes, No, Yes), since $\omega_2 \in P(\omega_1)$ but $P(\omega_1) \neq P(\omega_2)$.

69.2 (*Remembering numbers*) The set of states Ω is the set of integers and $P(\omega) = \{\omega - 1, \omega, \omega + 1\}$ for each $\omega \in \Omega$. The function P is not partitional: $1 \in P(0)$, for example, but $P(1) \neq P(0)$.

71.1 (*Information functions and knowledge functions*)

a. $P'(\omega)$ is the intersection of all events E for which $\omega \in K(E)$ and thus is the intersection of all E for which $P(\omega) \subseteq E$, and this intersection is $P(\omega)$ itself.

b. $K'(E)$ consists of all ω for which $P(\omega) \subseteq E$, where $P(\omega)$ is equal to the intersection of the events F that satisfy $\omega \in K(F)$. By K1, $P(\omega) \subseteq \Omega$.

Now, if $\omega \in K(E)$ then $P(\omega) \subseteq E$ and therefore $\omega \in K'(E)$. On the other hand if $\omega \in K'(E)$ then $P(\omega) \subseteq E$, or $E \supseteq \cap\{F \subseteq \Omega\colon K(F) \ni \omega\}$. Thus by K2 we have $K(E) \supseteq K(\cap\{F \subseteq \Omega\colon K(F) \ni \omega\})$, which by K3 is equal to $\cap\{K(F)\colon F \subseteq \Omega \text{ and } K(F) \ni \omega\})$, so that $\omega \in K(E)$. Hence $K(E) = K'(E)$.

71.2 (*Decisions and information*) Let a be the best act under P and let a' be the best act under P'. Then a' is feasible under P and the expected payoff from a' is

$$\sum_k \pi(P^k) \mathrm{E}_{\pi^k} u(a'(P'(P^k)), \omega),$$

where $\{P^1, \ldots, P^K\}$ is the partition induced by P, π^k is π conditional of P^k, $P'(P^k)$ is the member of the partition induced by P' that contains P^k, and we write $a'(P'(P^k))$ for the action $a'(\omega)$ for any $\omega \in P'(P^k)$. The result follows

from the fact that for each value of k we have

$$\mathrm{E}_{\pi^k} u(a(P^k), \omega) \geq \mathrm{E}_{\pi^k} u(a'(P'(P^k)), \omega).$$

76.1 (*Common knowledge and different beliefs*) Let $\Omega = \{\omega_1, \omega_2\}$, suppose that the partition induced by individual 1's information function is $\{\{\omega_1, \omega_2\}\}$ and that induced by individual 2's is $\{\{\omega_1\}, \{\omega_2\}\}$, assume that each individual's prior is $(\frac{1}{2}, \frac{1}{2})$, and let E be the event $\{\omega_1\}$. The event "individual 1 and individual 2 assign different probabilities to E" is $\{\omega \in \Omega: \rho(E|P_1(\omega)) \neq \rho(E|P_2(\omega))\} = \{\omega_1, \omega_2\}$, which is clearly self-evident, and hence is common knowledge in either state.

The proof of the second part follows the lines of the proof of Proposition 75.1. The event "the probability assigned by individual 1 to X exceeds that assigned by individual 2" is $E = \{\omega \in \Omega: \rho(X|P_1(\omega)) > \rho(X|P_2(\omega))\}$. If this event is common knowledge in the state ω then there is a self-evident event $F \ni \omega$ that is a subset of E and is a union of members of the information partitions of both individuals. Now, for all $\omega \in F$ we have $\rho(X|P_1(\omega)) > \rho(X|P_2(\omega))$, so that

$$\sum_{\omega \in F} \rho(\omega)\rho(X|P_1(\omega)) > \sum_{\omega \in F} \rho(\omega)\rho(X|P_2(\omega)).$$

But since F is a union of members of each individual's information partition both sides of this inequality are equal to $\rho(X \cap F)$, a contradiction. Hence E is not common knowledge.

76.2 (*Common knowledge and beliefs about lotteries*) Denote the value of the lottery in state ω by $L(\omega)$. Define the event E by

$$E = \{\omega \in \Omega: e_1(L|P_1(\omega)) > \eta \text{ and } e_2(L|P_2(\omega)) < \eta\},$$

where $e_i(L|P_i(\omega)) = \sum_{\omega' \in P_i(\omega)} \rho(\omega'|P_i(\omega))L(\omega')$ is individual i's belief about the expectation of the lottery. If this event is common knowledge in some state then there is a self-evident event $F \subseteq E$. Hence in every member of individual 1's information partition that is a subset of F the expected value of L exceeds η. Therefore $e_1(L|F) > \eta$: the expected value of the lottery given F is at least η. Analogously, the expected value of L given F is less than η, a contradiction.

Comment If this result were not true then a mutually profitable trade between the individuals could be made. The existence of such a pair of beliefs

is necessary for the existence of a rational expectations equilibrium in which the individuals are aware of the existing price, take it into consideration, and trade the lottery L even though they are risk-neutral.

Example for non-partitional information functions: Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\rho(\omega_i) = \frac{1}{3}$ for all $\omega \in \Omega$, $P_1(\omega) = \{\omega_1, \omega_2, \omega_3\}$ for all $\omega \in \Omega$, $P_2(\omega_1) = \{\omega_1, \omega_2\}$, $P_2(\omega_2) = \{\omega_2\}$, and $P_2(\omega_3) = \{\omega_2, \omega_3\}$ (so that P_2 is not partitional). Let $L(\omega_2) = 1$ and $L(\omega_1) = L(\omega_3) = 0$ and let $\eta = 0.4$. Then for all $\omega \in \Omega$ it is common knowledge that player 1 believes that the expectation of L is $\frac{1}{3}$ and that player 2 believes that the expectation of L is either 0.5 or 1.

81.1 (*Knowledge and correlated equilibrium*) By the rationality of player i in every state, for every $\omega \in \Omega$ the action $a_i(\omega)$ is a best response to player i's belief, which by assumption is derived from the common prior ρ and $P_i(\omega)$. Thus for all $\omega \in \Omega$ and all $i \in N$ the action $a_i(\omega)$ is a best response to the conditional probability derived from ρ, as required by the definition of correlated equilibrium.

6 Extensive Games with Perfect Information

94.2 (*Extensive games with* 2×2 *strategic forms*) First suppose that $(a_1', a_2') \sim_i (a_1', a_2'')$ for $i = 1, 2$. Then G is the strategic form of the extensive game with perfect information in Figure 21.1 (with appropriate assumptions on the players' preferences). The other case is similar.

Now assume that G is the strategic form of an extensive game Γ with perfect information. Since each player has only two strategies in Γ, for each player there is a single history after which he makes a (non-degenerate) move. Suppose that player 1 moves first. Then player 2 can move after only one of player 1's actions, say a_1''. In this case player 1's action a_1' leads to a terminal history, so that the combination of a_1' and either of the strategies of player 2 leads to the same terminal history; thus $(a_1', a_2') \sim_i (a_1', a_2'')$ for $i = 1, 2$.

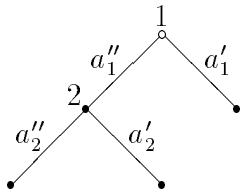

Figure 21.1 The game for the solution to Exercise 94.2.

98.1 (*SPE of Stackelberg game*) Consider the game in Figure 22.1. In this game (L, AD) is a subgame perfect equilibrium, with a payoff of $(1, 0)$, while the solution of the maximization problem is (R, C), with a payoff of $(2, 1)$.

99.1 (*Necessity of finite horizon for one deviation property*) In the (one-player) game in Figure 22.2 the strategy in which the player chooses d after every history satisfies the condition in Lemma 98.2 but is not a subgame perfect equilibrium.

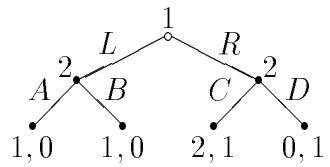

Figure 22.1 The extensive game in the solution of Exercise 98.1.

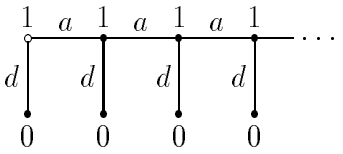

Figure 22.2 The beginning of a one-player infinite horizon game for which the one deviation property does not hold. The payoff to the (single) infinite history is 1.

100.1 (*Necessity of finiteness for Kuhn's theorem*) Consider the one-player game in which the player chooses a number in the interval $[0, 1)$, and prefers larger numbers to smaller ones. That is, consider the game $\langle \{1\}, \{\varnothing\} \cup [0, 1), P, \{\succsim_1\} \rangle$ in which $P(\varnothing) = 1$ and $x \succ_1 y$ if and only if $x > y$. This game has a finite horizon (the length of the longest history is 1) but has no subgame perfect equilibrium (since $[0, 1)$ has no maximal element).

In the infinite-horizon one-player game the beginning of which is shown in Figure 22.3 the single player chooses between two actions after every history. After any history of length k the player can choose to stop and obtain a payoff of $k + 1$ or to continue; the payoff if she continues for ever is 0. The game has no subgame perfect equilibrium.

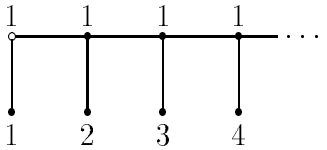

Figure 22.3 The beginning of a one-player game with no subgame perfect equilibrium. The payoff to the (single) infinite history is 0.

100.2 (*SPE of games satisfying no indifference condition*) The hypothesis is true for all subgames of length one. Assume the hypothesis for all subgames with length at most k. Consider a subgame $\Gamma(h)$ with $\ell(\Gamma(h)) = k+1$ and $P(h) = i$. For all actions a of player i such that $(h, a) \in H$ define $R(h, a)$ to be the outcome of some subgame perfect equilibrium of the subgame $\Gamma(h, a)$. By

hypothesis all subgame perfect equilibria outcomes of $\Gamma(h, a)$ are preference equivalent; in a subgame perfect equilibrium of $\Gamma(h)$ player i takes an action that maximizes $\succsim_i$ over $\{R(h, a): a \in A(h)\}$. Therefore player i is indifferent between any two subgame perfect equilibrium outcomes of $\Gamma(h)$; by the no indifference condition all players are indifferent among all subgame perfect equilibrium outcomes of $\Gamma(h)$.

We now show that the equilibria are interchangeable. For any subgame perfect equilibrium we can attach to every subgame the outcome according to the subgame perfect equilibrium if that subgame is reached. By the first part of the exercise the outcomes that we attach (or at least the rankings of these outcomes in the players' preferences) are independent of the subgame perfect equilibrium that we select. Thus by the one deviation property (Lemma 98.2), any strategy profile s'' in which for each player i the strategy s_i'' is equal to either s_i or s_i' is a subgame perfect equilibrium.

101.1 (*SPE and unreached subgames*) This follows directly from the definition of a subgame perfect equilibrium.

101.2 (*SPE and unchosen actions*) The result follows directly from the definition of a subgame perfect equilibrium.

101.3 (*Armies*) We model the situation as an extensive game in which at each history at which player i occupies the island and player j has at least two battalions left, player j has two choices: conquer the island or terminate the game. The first player to move is player 1. (We do not specify the game formally.)

We show that in every subgame in which army i is left with y_i battalions ($i = 1, 2$) and army j occupies the island, army i attacks if and only if either $y_i > y_j$, or $y_i = y_j$ and y_i is even.

The proof is by induction on $\min\{y_1, y_2\}$. The claim is clearly correct if $\min\{y_1, y_2\} \leq 1$. Now assume that we have proved the claim whenever $\min\{y_1, y_2\} \leq m$ for some $m \geq 1$. Suppose that $\min\{y_1, y_2\} = m + 1$. There are two cases.

- either $y_i > y_j$, or $y_i = y_j$ and y_i is even: If army i attacks then it occupies the island and is left with $y_i - 1$ battalions. By the induction hypothesis army j does not launch a counterattack in any subgame perfect equilibrium, so that the attack is worthwhile.

- either $y_i < y_j$, or $y_i = y_j$ and y_i is odd: If army i attacks then it occupies the island and is left with $y_i - 1$ battalions; army j is left with y_j battalions. Since either $y_i - 1 < y_j - 1$ or $y_i - 1 = y_j - 1$ and is even, it follows from the inductive hypothesis that in all subgame perfect equilibria there is a counterattack. Thus army i is better off not attacking.

Thus the claim is correct whenever $\min\{y_1, y_2\} \leq m + 1$, completing the inductive argument.

102.1 (*ODP and Kuhn's theorem with chance moves*)

One deviation property: The argument is the same as in the proof of Lemma 98.2.

Kuhn's theorem: The argument is the same as in the proof of Proposition 99.2 with the following addition. If $P(h^*) = c$ then $R(h^*)$ is the lottery in which $R(h^*, a)$ occurs with probability $f_c(a|h)$ for each $a \in A(h^*)$.

103.1 (*Three players sharing pie*) The game is given by

- $N = \{1, 2, 3\}$
- $H = \{\varnothing\} \cup X \cup \{(x, y)\colon x \in X \text{ and } y \in \{yes, no\} \times \{yes, no\}\}$ where $X = \{x \in \mathbb{R}^3_+ \colon \sum_{i=1}^3 x_i = 1\}$
- $P(\varnothing) = 1$ and $P(x) = \{2, 3\}$ if $x \in X$
- for each $i \in N$ we have $(x, (yes, yes)) \succ_i (z, (yes, yes))$ if and only if $x_i > z_i$; if $(A, B) \neq (yes, yes)$ then $(x, (yes, yes)) \succ_i (z, (A, B))$ if $x_i > 0$ and $(x, (yes, yes)) \sim_i (z, (A, B))$ if $x_i = 0$; if $(A, B) \neq (yes, yes)$ and $(C, D) \neq (yes, yes)$ then $(x, (C, D)) \sim_i (z, (A, B))$ for all $x \in X$ and $z \in X$.

In each subgame that follows a proposal x of player 1 there are two types of Nash equilibria. In one equilibrium, which we refer to as $Y(x)$, players 2 and 3 both accept x. In all the remaining equilibria the proposal x is not implemented; we refer to the set of these equilibria as $N(x)$. If both $x_2 > 0$ and $x_3 > 0$ then $N(x)$ consists of the single equilibrium in which players 2 and 3 both reject x. If $x_i = 0$ for either $i = 2$ or $i = 3$, or both, then $N(x)$ contains in addition equilibria in which a player who is offered 0 rejects the proposal and the other player accepts the proposal.

Consequently the equilibria of the entire game are the following.

- For any division x, player 1 proposes x. In the subgame that follows the proposal x of player 1, the equilibrium is $Y(x)$. In the subgame that follows any proposal y of player 1 in which $y_1 > x_1$, the equilibrium is in $N(y)$. In the subgame that follows any proposal y of player 1 in which $y_1 < x_1$, the equilibrium is either $Y(y)$ or is in $N(y)$.

- For any division x, player 1 proposes x. In the subgame that follows any proposal y of player 1 in which $y_1 > 0$, the equilibrium is in $N(y)$. In the subgame that follows any proposal y of player 1 in which $y_1 = 0$, the equilibrium is either $Y(y)$ or is in $N(y)$.

103.2 (*Naming numbers*) The game is given by

- $N = \{1, 2\}$
- $H = \{\varnothing\} \cup \{Stop, Continue\} \cup \{(Continue, y)\colon y \in Z \times Z\}$ where Z is the set of nonnegative integers
- $P(\varnothing) = 1$ and $P(Continue) = \{1, 2\}$
- the preference relation of each player is determined by the payoffs given in the question.

In the subgame that follows the history *Continue* there is a unique subgame perfect equilibrium, in which both players choose 0. Thus the game has a unique subgame perfect equilibrium, in which player 1 chooses *Stop* and, if she chooses *Continue*, both players choose 0.

Note that if the set of actions of each player after player 1 chooses *Continue* were bounded by some number M then there would be an additional subgame perfect equilibrium in which player 1 chooses *Continue* and each player names M, with the payoff profile (M^2, M^2).

103.3 (*ODP and Kuhn's theorem with simultaneous moves*)

One deviation property: The argument is the same as in the proof of Lemma 98.2.

Kuhn's theorem: Consider the following game (which captures the same situation as Matching Pennies (Figure 17.3)):

- $N = \{1, 2\}$
- $H = \{\varnothing\} \cup \{x \in \{\text{Head}, \text{Tail}\} \times \{\text{Head}, \text{Tail}\}$

- $P(\emptyset) = \{1, 2\}$
- (Head, Head) $\sim_1$ (Tail, Tail) $\succ_1$ (Head, Tail) $\sim_1$ (Tail, Head) and (Head, Tail) $\sim_2$ (Tail, Head) $\succ_2$ (Head, Head) $\sim_2$ (Tail, Tail).

This game has no subgame perfect equilibrium.

108.1 (ϵ-*equilibrium of centipede game*) Consider the following pair of strategies. In every period before k both players choose C; in every subsequent period both players choose S. The outcome is that the game stops in period k. We claim that if $T \geq 1/\epsilon$ then this strategy pair is a Nash equilibrium. For concreteness assume that k is even, so that it is player 2's turn to act in period k. Up to period $k-2$ both players are worse off if they choose S rather than C. In period $k-1$ player 1 gains $1/T \leq \epsilon$ by choosing S. In period k player 2 is better off choosing S (given the strategy of player 1), and in subsequent periods the action that each player chooses has no effect on the outcome. Thus the strategy pair is an ϵ-equilibrium of the game.

114.1 (*Variant of the game* Burning money) Player 1 has eight strategies, each of which can be written as (x, y, z), where $x \in \{0, D\}$ and y and z are each members of $\{B, S\}$, y being the action that player 1 plans in BoS if player 2 chooses 0 and z being the action that player 1 plans in BoS if player 2 chooses D. Player 2 has sixteen strategies, each of which can be written as a pair of members of the set $\{(0, B), (0, S), (D, B), (D, S)\}$, the first member of the pair being player 2's actions if player 1 chooses 0 and the second member of the pair being player 2's actions if player 1 chooses D.

Weakly dominated actions can be iteratively eliminated as follows.

1. (D, S, S) is weakly dominated for player 1 by $(0, B, B)$

   Every strategy (a, b) of player 2 in which either a or b is (D, B) is weakly dominated by the strategy that differs only in that (D, B) is replaced by $(0, S)$.

2. Every strategy (x, y, B) of player 1 is weakly dominated by (x, y, S) (since there is no remaining strategy of player 2 in which he chooses (D, B)).

3. Every strategy (a, b) of player 2 in which b is either $(0, B)$ or $(0, S)$ is weakly dominated by the strategy that differs only in that b is replaced by (D, S) (since in every remaining strategy player 1 chooses S after player 2 chooses D).

	$(0,B),(D,S))$	$((0,S),(D,S))$	$((D,S),(D,S))$
$(0,B,S)$	$3,1$	$0,0$	$1,2$
$(0,S,S)$	$0,0$	$1,3$	$1,2$
(D,B,S)	$0,2$	$0,2$	$0,2$

Figure 27.1 The game in Exercise 114.1 after three rounds of elimination of weakly dominated strategies.

	AA	AB	BA	BB
$0A$	$2,2$	$2,2$	$0,0$	$0,0$
$0B$	$0,0$	$0,0$	$1,1$	$1,1$
DA	$1,2$	$-1,0$	$1,2$	$-1,0$
DB	$-1,0$	$0,1$	$-1,0$	$0,1$

Figure 27.2 The game for Exercise 114.2.

The game that remains is shown in Figure 27.1.

4. (D,B,S) is weakly dominated for player 1 by $(0,B,S)$
   $(0,B),(D,S))$ is weakly dominated for player 2 by $((D,S),(D,S))$
5. $(0,B,S)$ is weakly dominated for player 1 by $(0,S,S)$
6. $((D,S),(D,S))$ is strictly dominated for player 2 by $((0,S),(D,S))$

The only remaining strategy pair is $((0,S,S),((0,S),(D,S)))$, yielding the outcome $(1,3)$ (the one that player 2 most prefers).

114.2 (*Variant of the game* Burning money) The strategic form of the game is given in Figure 27.2. Weakly dominated actions can be eliminated iteratively as follows.

1. DB is weakly dominated for player 1 by $0B$
2. AB is weakly dominated for player 2 by AA
   BB is weakly dominated for player 2 by BA
3. $0B$ is strictly dominated for player 1 by DA
4. BA is weakly dominated for player 2 by AA
5. DA is strictly dominated for player 1 by $0A$

The single strategy pair that remains is $(0A, AA)$.

7 A Model of Bargaining

123.1 (*One deviation property for bargaining game*) The proof is similar to that of Lemma 98.2; the sole difference is that the existence of a profitable deviant strategy that differs from s^* after a finite number of histories follows from the fact that the single infinite history is the worst possible history in the game.

125.2 (*Constant cost of bargaining*)

a. It is straightforward to check that the strategy pair is a subgame perfect equilibrium. Let $M_i(G_i)$ and $m_i(G_i)$ be as in the proof of Proposition 122.1 for $i = 1, 2$. By the argument for (124.1) with the roles of the players reversed we have $M_2(G_2) \leq 1 - m_1(G_1) + c_1$, or $m_1(G_1) \leq 1 - M_2(G_2) + c_1$. Now suppose that $M_2(G_2) \geq c_2$. Then by the argument for (123.2) with the roles of the players reversed we have $m_1(G_1) \geq 1 - M_2(G_2) + c_2$, a contradiction (since $c_1 < c_2$). Thus $M_2(G_2) < c_2$. But now the argument for (123.2) implies that $m_1(G_1) \geq 1$, so that $m_1(G_1) = 1$ and hence $M_1(G_1) = 1$. Since (124.1) implies that $M_2(G_2) \leq 1 - m_1(G_1) + c_1$ we have $M_2(G_2) \leq c_1$; by (123.2) we have $m_2(G_2) \geq c_1$, so that $M_2(G_2) = m_2(G_2) = c_1$. The remainder of the argument follows as in the proof of Proposition 122.1.

b. First note that for any pair (x^*, y^*) of proposals in which $x_1^* \geq c$ and $y_1^* = x_1^* - c$ the pair of strategies described in Proposition 122.1 is a subgame perfect equilibrium. Refer to this equilibrium as $E(x^*)$.

Now suppose that $c < \frac{1}{3}$. An example of an equilibrium in which agreement is reached with delay is the following. Player 1 begins by proposing $(1, 0)$. Player 2 rejects this proposal, and play continues as in the equilibrium $E(\frac{1}{3}, \frac{2}{3})$. Player 2 rejects also any proposal x in which $x_1 > c$ and accepts all other proposals; in each of these cases play continues as in the equilibrium $E(c, 1-c)$. An interpretation of this equilibrium is that player 2 regards player 1's making a proposal different from $(1, 0)$ as a sign of "weakness".

127.1 (*One-sided offers*) We argue that the following strategy pair is the unique subgame perfect equilibrium: player 1 always proposes $b(1)$ and player 2 always accepts all offers. It is clear that this is a subgame perfect equilibrium. To show that it is the only subgame perfect equilibrium choose $\delta \in (0,1)$ and suppose that player i's preferences are represented by the function $\delta^t u_i(x)$ with $u_j(b(i)) = 0$. Let M_2 be the supremum of player 2's payoff and let m_1 be the infimum of player 1's payoff in subgame perfect equilibria of the game. (Note that the definitions of M_2 and m_1 differ from those in the proof of Proposition 122.1.) Then $m_1 \geq \phi(\delta M_2)$ (by the argument for (123.2) in the proof of Proposition 122.1) and $m_1 \leq \phi(M_2)$. Hence $M_2 \leq \delta M_2$, so that $M_2 = 0$ and hence the agreement reached is $b(1)$, and this must be reached immediately.

128.1 (*Finite grid of possible offers*) *a*. For each player i let σ_i be the strategy in which player i always proposes x and accepts a proposal y if and only if $y_i \geq x_i$ and let $\delta \geq 1 - \epsilon$. The outcome of (σ_1, σ_2) is $(x, 0)$. To show that (σ_1, σ_2) is a subgame perfect equilibrium the only significant step is to show that it is optimal for each player i to reject the proposal in which he receives $x_i - \epsilon$. If he does so then his payoff is δx_i, so that we need $\delta x_i \geq x_i - \epsilon$, or $\delta \geq 1 - \epsilon/x_i$, which is guaranteed by our choice of $\delta \geq 1 - \epsilon$.

b. Let $(x^*, t^*) \in X \times T$ (the argument for the outcome D is similar). For $i = 1, 2$, define the strategy $\overline{\sigma}_i$ as follows.

- in any period $t < t^*$ at which no player has previously deviated, propose b^i (the best agreement for player i) and reject any proposal other than b^i
- if any period $t \geq t^*$ at which no player has previously deviated, propose x^* and accept a proposal y if and only if $y \succsim_i x^*$.
- in any period at which some player has previously deviated, follow the equilibrium defined in part *a* for $x = (0,1)$ if player 1 was the first to have deviated and for $x = (1,0)$ if player 2 was the first to have deviated.

The outcome of the strategy pair $(\overline{\sigma}_1, \overline{\sigma}_2)$ is (x^*, t^*). If $\delta \geq 1 - \epsilon$ the strategy pair is a subgame perfect equilibrium. Given part *a*, the significant step is to show that neither player wants to deviate through period t^*, which is the case since any deviation that does not end the game leads to an outcome in which the deviator gets 0, and any unplanned acceptance is of a proposal that gives the responder 0.

c. First we show that $\Gamma(\epsilon)$ has a subgame perfect equilibrium for every value of ϵ. For any real number x, denote by $[x]$ the smallest integral multiple of ϵ that is at least equal to x. Let $z = [1/(1+\delta)] - \epsilon$ and $z' = [1/(1+\delta)]$. There are two cases.

- If $z \geq (1-\epsilon)/(1+\delta)$ then $\Gamma(\epsilon)$ has a subgame perfect equilibrium in which the players' strategies have the same structure as those in Proposition 122.1, with $x^* = (z, 1-z)$ and $y^* = (1-z, z)$. It is straightforward to show that this strategy pair is a subgame perfect equilibrium (in particular, it is optimal for a responder to accept an offer in which his payoff is $1-z$ and reject an offer in which his payoff is $1-z-\epsilon$).

- If $z < (1-\epsilon)/(1+\delta)$ then $\Gamma(\epsilon)$ has a subgame perfect equilibrium in which each player uses the same strategy, which has two "states": state z, in which the proposal gives the proposer a payoff of z and an offer is accepted if and only if the responder's payoff is at least $1-z$, and state z', in which the proposal gives the proposer a payoff of z' and an offer is accepted if and only if the responder's payoff is at least $1-z'$. Initially both players' strategies are in state z; subsequently any deviation in one of the states triggers a switch to the other state. It is straightforward to check that in state z a responder should accept $(z, 1-z)$ and reject $(z+\epsilon, 1-z-\epsilon)$ and in state z' a responder should accept $(z', 1-z')$ and reject $(z'+\epsilon, 1-z'-\epsilon)$.

Now let M be the supremum of a player's payoff over the subgame perfect equilibria of subgames in which he makes the first proposal; let m be the corresponding infimum. By the arguments for (123.2) and (124.1) we have $m \geq 1-[\delta M]$ and $1-\delta m \geq M$, from which it follows that $m \geq 1/(1+\delta) - \epsilon/(1-\delta^2)$ and $M \leq 1/(1+\delta) + \delta\epsilon/(1-\delta^2)$. Thus player 1's payoff in any subgame perfect equilibrium is close to $1/(1+\delta)$ when ϵ is small. Since player 2 can reject any proposal of player 1 and become a proposer, his subgame perfect equilibrium payoff is at least δm; since player 1's payoff is at least m, player 2's payoff is at most $1-m$. If follows that player 2's payoff in any subgame perfect equilibrium is close to $\delta/(1+\delta)$ when ϵ is small. This is, the difference between each player's payoff in every subgame perfect equilibrium of $\Gamma(\epsilon)$ and his payoff in the unique subgame perfect equilibrium of $\Gamma(0)$ can be made arbitrarily small by decreasing ϵ.

Finally, the proposer's payoff in any subgame perfect equilibrium is at least m and the responder's payoff is at least δm, and by the inequality for m above we have $m + \delta m \geq 1 - \epsilon/(1-\delta)$, so that the sum of the players' payoffs in any

subgame perfect equilibrium exceeds δ if ϵ is small enough. Thus for ϵ small enough agreement is reached immediately in any subgame perfect equilibrium.

129.1 (*Outside options*) It is straightforward to check that the strategy pair described is a subgame perfect equilibrium. The following proof of uniqueness is taken from Osborne and Rubinstein (1990).

Let M_1 and M_2 be the suprema of player 1's and player 2's payoffs over subgame perfect equilibria of the subgames in which players 1 and 2, respectively, make the first offer. Similarly, let m_1 and m_2 be the infima of these payoffs. Note that $(Out, 0) \precsim_2 (y^*, 1)$ if and only if $b \leq \delta/(1+\delta)$. We proceed in a number of steps.

Step 1. $m_2 \geq 1 - \delta M_1$.

The proof is the same as that for (123.2) in the proof of Proposition 122.1.

Step 2. $M_1 \leq 1 - \max\{b, \delta m_2\}$.

Proof. Since Player 2 obtains the payoff b by opting out, we must have $M_1 \leq 1 - b$. The fact that $M_1 \leq 1 - \delta m_2$ follows from the same argument as for (124.1) in the proof of Proposition 122.1.

Step 3. $m_1 \geq 1 - \max\{b, \delta M_2\}$ and $M_2 \leq 1 - \delta m_1$.

The proof is analogous to those for Steps 1 and 2.

Step 4. If $\delta/(1+\delta) \geq b$ then $m_i \leq 1/(1+\delta) \leq M_i$ for $i = 1, 2$.

Proof. These inequalities follow from the fact that in the subgame perfect equilibrium described in the text player 1 obtains the payoff $1/(1+\delta)$ in any subgame in which she makes the first offer, and player 2 obtains the same utility in any subgame in which he makes the first offer.

Step 5. If $\delta/(1+\delta) \geq b$ then $M_1 = m_1 = 1/(1+\delta)$ and $M_2 = m_2 = 1/(1+\delta)$.

Proof. By Step 2 we have $1 - M_1 \geq \delta m_2$, and by Step 1 we have $m_2 \geq 1 - \delta M_1$, so that $1 - M_1 \geq \delta - \delta^2 M_1$, and hence $M_1 \leq 1/(1+\delta)$. Hence $M_1 = 1/(1+\delta)$ by Step 4.

Now, by Step 1 we have $m_2 \geq 1 - \delta M_1 = 1/(1+\delta)$. Hence $m_2 = 1/(1+\delta)$ by Step 4.

Again using Step 4 we have $\delta M_2 \geq \delta/(1+\delta) \geq b$, and hence by Step 3 we have $m_1 \geq 1 - \delta M_2 \geq 1 - \delta(1 - \delta m_1)$. Thus $m_1 \geq 1/(1+\delta)$. Hence $m_1 = 1/(1+\delta)$ by Step 4.

Finally, by Step 3 we have $M_2 \leq 1 - \delta m_1 = 1/(1+\delta)$, so that $M_2 = 1/(1+\delta)$ by Step 4.

Step 6. If $b \geq \delta/(1+\delta)$ then $m_1 \leq 1-b \leq M_1$ and $m_2 \leq 1-\delta(1-b) \leq M_2$.

Proof. These inequalities follow from the subgame perfect equilibrium described in the text (as in Step 4).

Step 7. If $b \geq \delta/(1+\delta)$ then $M_1 = m_1 = 1-b$ and $M_2 = m_2 = 1-\delta(1-b)$.

Proof. By Step 2 we have $M_1 \leq 1-b$, so that $M_1 = 1-b$ by Step 6. By Step 1 we have $m_2 \geq 1-\delta M_1 = 1-\delta(1-b)$, so that $m_2 = 1-\delta(1-b)$ by Step 6.

Now we show that $\delta M_2 \leq b$. If $\delta M_2 > b$ then by Step 3 we have $M_2 \leq 1-\delta m_1 \leq 1-\delta(1-\delta M_2)$, so that $M_2 \leq 1/(1+\delta)$. Hence $b < \delta M_2 \leq \delta/(1+\delta)$, contradicting our assumption that $b \geq \delta/(1+\delta)$.

Given that $\delta M_2 \leq b$ we have $m_1 \geq 1-b$ by Step 3, so that $m_1 = 1-b$ by Step 6. Further, $M_2 \leq 1-\delta m_1 = 1-\delta(1-b)$ by Step 3, so that $M_2 = 1-\delta(1-b)$ by Step 6.

Thus in each case the subgame perfect equilibrium outcome is unique. The argument that the subgame perfect equilibrium strategies are unique is the same as in the proof of Proposition 122.1.

130.2 (*Risk of breakdown*) The argument that the strategy pair is a subgame perfect equilibrium is straightforward. The argument for uniqueness is analogous to that in Proposition 122.1, with $1-\alpha$ playing the role of δ_i for $i = 1, 2$.

131.1 (*Three-player bargaining*) First we argue that in any subgame perfect equilibrium the offer made by each player is immediately accepted. For $i = 1, 2, 3$, let U^i be the equilibrium payoff profile in the subgames beginning with offers by player i. (Since the strategies are stationary these profiles are independent of history.) If player 1 proposes an agreement in which each of the other player's payoffs exceeds δU_j^2 then those players must both accept. Thus player 1's equilibrium payoff U_1^1 is at least $1-\delta U_2^2 - U_3^2$. In any equilibrium in which player 1's offer is rejected her payoff is at most $\delta(1-U_2^2-U_3^2) < 1-\delta U_2^2 - U_3^2$, so that in any equilibrium player 1's offer is accepted. Similarly the offers of player 2 and player 3 are accepted immediately.

Now, let the proposals made by the three players be x^*, y^*, and z^*. Then the requirement that player 1's equilibrium proposal be optimal implies that $x_2^* = \delta y_2^*$ and $x_3^* = \delta y_3^*$; similarly $y_1^* = \delta z_1^*$ and $y_3^* = \delta z_3^*$, and $z_1^* = \delta x_1^*$ and $z_2^* = \delta x_2^*$. The unique solution of these equations yields the offer x^* described in the problem.

8 Repeated Games

139.1 (*Discount factors that differ*) Consider a two-player game in which the constituent game has two payoff profiles, $(1,0)$ and $(0,1)$. Let (v^t) be the sequence of payoff profiles of the constituent game in which $v^1 = (0,1)$ and $v^t = (1,0)$ for all $t \geq 2$. The payoff profile associated with this sequence is $(\delta_1, 1-\delta_2)$. Whenever $\delta_1 \neq \delta_2$ this payoff profile is not feasible. In particular, when δ_1 is close to 1 and δ_2 is close to 0 the payoff profile is close to $(1,1)$, which Pareto dominates all feasible payoff profiles of the constituent game.

143.1 (*Strategies and finite machines*) Consider the strategy of player 1 in which she chooses C then D, followed by C and two D's, followed by C and three D's, and so on, independently of the other players' behavior. Since there is no cycle in this sequence, the strategy cannot be executed by a machine with finitely many states.

144.2 (*Machine that guarantees v_i*) Let player 2's machine be $\langle Q_2, q_2^0, f_2, \tau_2 \rangle$; a machine that induces a payoff for player 1 of at least v_1 is $\langle Q_1, q_1^0, f_1, \tau_1 \rangle$ where

- $Q_1 = Q_2$.
- $q_1^0 = q_2^0$.
- $f_1(q) = b_1(f_2(q))$ for all $q \in Q_2$.
- $\tau_1(q,a) = \tau_2(q,a)$ for all $q \in Q_2$ and $a \in A$.

This machine keeps track of player 2's state and always responds to player 2's action in such a way that it obtains a payoff of at least v_1.

145.1 (*Machine for Nash folk theorem*) Let $N = \{1, \ldots, n\}$. A machine that executes s_i is $\langle Q_i, q_i^0, f_i, \tau_i \rangle$ where

- $Q_i = \{S_1, \ldots, S_\gamma, P_1, \ldots, P_n\}$.
- $q_i^0 = S_1$.
- $f_i(q) = \begin{cases} a_i^\ell & \text{if } q = S_\ell \text{ or } q = P_i \\ (p_{-j})_i & \text{if } q = P_j \text{ for } i \neq j. \end{cases}$
- $\tau_i(S_\ell, a) = \begin{cases} P_j & \text{if } a_j \neq a_j^\ell \text{ and } a_i = a_i^\ell \text{ for all } i \neq j \\ S_{\ell+1 \,(\mathrm{mod}\,\gamma)} & \text{otherwise} \end{cases}$

  and $\tau_i(P_j, a) = P_j$ for all $a \in A$.

146.1 (*Example with discounting*) We have $(v_1, v_2) = (1, 1)$, so that the payoff of player 1 in every subgame perfect equilibrium is at least 1. Since player 2's payoff always exceeds player 1's payoff we conclude that player 2's payoff in any subgame perfect equilibria exceeds 1. The path $((A, A), (A, A), \ldots)$ is not a subgame perfect equilibrium outcome path since player 2 can deviate to D, achieving a payoff of 5 in the first period and more than 1 in the subsequent subgame, which is better for him than the constant sequence $(3, 3, \ldots)$.

Comment We use only the fact that player 2's discount factor is at most $\frac{1}{2}$.

148.1 (*Long- and short-lived players*) First note that in any subgame perfect equilibrium of the game, the action taken by the opponent of player 1 in any period t is a one-shot best response to player 1's action in period t.

a. The game has a unique subgame perfect equilibrium, in which player 1 chooses D in every period and each of the other players chooses D.

b. Choose a sequence of outcomes (C, C) and (D, D) whose average payoff to player 1 is x. Player 1's strategy makes choices consistent with this path so long as the previous outcomes were consistent with the path; subsequent to any deviation it chooses D for ever. Her opponent's strategy in any period t makes the choice consistent with the path so long as the previous outcomes were consistent with the path, and otherwise chooses D.

152.1 (*Game that is not full dimensional*)

a. For each $i \in N$ we have $v_i = 0$ (if one of the other players chooses 0 and the other chooses 1 then player i's payoff is 0 regardless of his action) and

the maximum payoff of every player is 1. Thus the set of enforceable payoff profiles is $\{(w_1, w_2, w_3): w_i \in [0,1] \text{ for } i = 1,2,3\}$.

b. Let m be the minimum payoff of any player in a subgame perfect equilibria of the repeated game. Consider a subgame perfect equilibrium in which every player's payoff is m; let a^1 be the action profile chosen by the players in the first period in this subgame perfect equilibrium. Then for some player i we have either $a_j^1 \leq \frac{1}{2}$ and $a_k^1 \leq \frac{1}{2}$ or $a_j^1 \geq \frac{1}{2}$ and $a_k^1 \geq \frac{1}{2}$ where j and k are the players other than i. Thus by deviating from a_i^1 player i can obtain at least $\frac{1}{4}$ in period 1; subsequently he obtains at least $\delta m/(1-\delta)$. Thus in order for the deviation to be unprofitable we require $\frac{1}{4} + \delta m/(1-\delta) \leq m/(1-\delta)$ or $m \geq \frac{1}{4}$.

c. The full dimensionality assumption in Proposition 151.1 (on the collection $\{a(i)\}_{i \in N}$ of strictly enforceable outcomes) is violated by the game G: for any outcomes $a(1)$ and $a(2)$, if $a(1) \succ_2 a(2)$ then also $a(1) \succ_1 a(2)$.

153.2 (*One deviation property for discounted repeated game*) Let $s = (s_i)_{i \in N}$ be a strategy profile in the repeated game and let $(v^t)_{t=1}^{\infty}$ be the infinite sequence of payoff profiles of G that s induces; let $U_i(s) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v_i^t$, player i's payoff in the repeated game when the players adopt the strategy profile s. For any history $h = (a^1, \ldots, a^t)$ let

$$W_i(s,h) = (1-\delta)\sum_{k=1}^{\infty} \delta^{k-1} u_i(a^{t+k}),$$

where $(a^{t+k})_{k=1}^{\infty}$ is the sequence of action profiles that s generates after the history h. That is, $W_i(s,h)$ is player i's payoff, discounted to period $t+1$, in the subgame starting after the history h when the players use the strategy profile s.

If a player can gain by a one-period deviation then the strategy profile is obviously not a subgame perfect equilibrium.

Now assume that no player can gain by a one-period deviation from s after any history but there is a history h after which player i can gain by switching to the strategy s_i'. For concreteness assume that h is the empty history, so that $U_i(s_{-i}, s_i') > U_i(s)$. Given that the players' preferences are represented by the discounting criterion, for every $\epsilon > 0$ there is some period T such that any change in player i's payoffs in any period after T does not change player i's payoff in the repeated game by more than ϵ. Thus we can assume that there exists some period T such that s_i' differs from s_i only in the first T periods. For any positive integer t let $h^t = (a^1, \ldots, a^t)$ be the sequence of outcomes of G induced by (s_{-i}, s_i') in the first t periods of the repeated game. Then since

s_i and s_i' differ only in the first T periods we have

$$U_i(s_{-i}, s_i') = (1-\delta)\sum_{k=1}^{T} \delta^{k-1} u_i(a^k) + \delta^T W_i(s, h^T).$$

Now, since no player can gain by deviating in a single period after any history, player i cannot gain by deviating from s_i in the first period of the subgame that follows the history h^{T-1}. Thus $(1-\delta)u_i(a^T) + \delta W_i(s, h^T) \leq W_i(s, h^{T-1})$ and hence

$$U_i(s_{-i}, s_i') \leq (1-\delta)\sum_{k=1}^{T-1} \delta^{k-1} u_i(a^k) + \delta^{T-1} W_i(s, h^{T-1}).$$

Continuing to work backwards period by period leads to the conclusion that

$$U_i(s_{-i}, s_i') \leq W_i(s, \varnothing) = U_i(s),$$

contradicting our assumption that player i's strategy s_i' is a profitable deviation.

157.1 (*Nash folk theorem for finitely repeated games*) For each $i \in N$ let $\hat{a}^i$ be a Nash equilibrium of G in which player i's payoff exceeds his minmax payoff v_i. To cover this case, the strategy in the proof of Proposition 156.1 needs to be modified as follows.

- The single state *Nash* is replaced by a collection of states $Nash^i$ for $i \in N$.
- In $Nash^i$ each player j chooses the action $\hat{a}_j^i$.
- The transition from $Norm^{T-L}$ is to $Nash^1$, and the transition from $Nash^k$ is to $Nash^{k+1 \pmod{|N|}}$
- $L = K|N|$ for some integer K and K is chosen to be large enough that $\max_{a_i \in A_i} u_i(a_{-i}^*, a_i) - u_i(a^*) \leq K\left(\sum_{j \in N} u_i(\hat{a}^j) - |N| v_i\right)$ for all $i \in N$.
- T^* is chosen so that $|[(T^* - L)u_i(a^*) + K\sum_{j \in N} u_i(\hat{a}^j)]/T^* - u_i(a^*)| < \epsilon$.

9 Complexity Considerations in Repeated Games

169.1 (*Unequal numbers of states in machines*) Consider the game $\langle\{1,2,3\},\{A_i\},\{u_i\}\rangle$ in which $A_1 = A_2 \times A_3$, $A_2 = \{\alpha,\beta\}$, $A_3 = \{x,y,z\}$, and $u_1(a) = 1$ if $a_1 = (a_2,a_3)$, $u_i(a) = 1$ if $a_i = (a_1)_{i-1}$ for $i = 2$, 3, and all other payoffs are 0. Suppose that player 2 uses a machine with a cycle of length 2, player 3 uses a machine with a cycle of length 3, and player 1 wants to coordinate with players 2 and 3. Then player 1 needs to have six states in her machine. Precisely, let $M_1 = \langle Q_1, q_1^0, f_1, \tau_1\rangle$ where $Q_1 = A_1$, $q_1^0 = (\alpha,x)$, $f_1(q) = q$ for all $q \in Q_1$, and for all $a \in A$ the state $\tau_1(q,a)$ is that which follows q in the sequence consisting of repetitions of the cycle (α,x), (β,y), (α,z), (β,x), (α,y), (β,z). Define M_2 as cycling between α and β and M_3 as cycling between x, y, and z. Then (M_1,M_2,M_3) is a Nash equilibrium of the machine game.

173.1 (*Equilibria of the* Prisoner's Dilemma)

a. It is easy to see that neither player can increase his payoff in the repeated game by using a different machine: every deviation initiates a sequence of four periods in which the other player chooses D, more than wiping out the immediate gain to the deviation if δ is close enough to 1. To show that a player cannot obtain the same payoff in the repeated game by a less complex machine assume that player 1 uses a machine M_1 with fewer than five states and player 2 uses the machine M. The pair (M_1,M) generates a cycle in which either R_2 is not reached and thus the average is less than 1, or R_2 is reached when player 1 plays D and is followed by at least four periods in which player 2 plays D, yielding a discounted average payoff close to $(1+1+1+1+5)/5 = 9/5$ when δ is close to 1. Thus (M,M) is a Nash equilibrium of the machine game.

b. The new pair of machines is not a Nash equilibrium since a player can obtain the same payoff by omitting the state I_3 and transiting from I_2 to R_2 if the other player chooses D.

173.2 (*Equilibria with introductory phases*) First note that in every equilibrium in which (C, C) is one of the outcomes on the equilibrium path the set of outcomes on the path is either $\{(C, C)\}$ or $\{(C, C), (D, D)\}$.

Now suppose that there is an equilibrium that has no introductory phase. Denote the states in the cycle by $q^1, \ldots, q^K$ and the equilibrium payoff of each player by z. Suppose that in state q^k the outcome is (C, C). Then a deviation to D by player 1 in state q^k must be deterred: suppose that in response to such a deviation player 2's machine goes to state q^m. It follows that player 1's average payoff from state q^{k+1} through q^{m-1} exceeds z, since if it were not then her average payoff in states q^m through q^k (where we take q^1 to be the state that follows q^K) would be at least z, so that a deviation in state q^k would be profitable. We conclude that there exists some k' such that player 1's payoff in states q^{k+1} through $q^{k'-1}$ exceeds z; without loss of generality we can assume that the outcome in state $q^{k'}$ is (C, C).

Now repeat the procedure starting from the state $q^{k'}$. Again we conclude that there exists some k'' such that player 1's payoff in states $q^{k'+1}$ through $q^{k''-1}$ exceeds z and the outcome in state $q^{k''}$ is (C, C). If we continue in the same manner then, since K is finite, we eventually return to the state q^k that we began with. In this way we cover the cycle an integer number of times and thus conclude that the average payoff in the cycle $q^1, \ldots, q^K$ exceeds z, contrary to our original assumption.

174.1 (*Case in which constituent game is extensive game*)

a. From Lemma 170.1 the set of outcomes that occurs in an equilibrium path is either a subset of $\{(A, B), (B, A)\}$ or a subset of $\{(A, A), (B, B)\}$. The former case is impossible by the following argument. The path in which the outcome in every period is (B, A) is not an equilibrium outcome since players 1 and 2 then use one-state machines that play B and A respectively, and player 1 can profitably gain by switching to the one-state machine that plays A. Every other path that contains both the outcomes (A, B) and (B, A) cannot be an equilibrium path since player 1's payoff is less than 2, which he can achieve in every period by using a one-state machine that always plays B. The remaining possibilities are that the outcome is (B, B) in every period or that it is either (A, A) or (B, B).

b. A Nash equilibrium can be constructed by having a long enough introductory phase in which (B, B) occurs in every period, with deviations in the cycling phase sending each machine back to its initial state.

c. Any Nash equilibrium of the machine game for the repeated extensive

game is a Nash equilibrium of the machine game for the repeated strategic game. Thus by part (a) in all possible equilibria of the machine game for the repeated extensive game the outcome is either (A, A) or (B, B) in every period. But if there is any occurrence of (A, A) then player 2 can drop the state in which he chooses B and simply choose A in every period. (If player 1 chooses B then she does not observe player 2's choice, so that this change in player 2's machine does not affect the equilibrium path.) Thus in the only possible equilibria the outcome is (B, B) in every period; it is clear that both players choosing a one-state machine that chooses B in every period is indeed a Nash equilibrium.

10 Implementation Theory

182.1 (*DSE-implementation with strict preferences*) Given Lemma 181.4 we need to show only that if a choice function is truthfully DSE-implementable then it is DSE-implementable. Suppose that the choice function $f: \mathcal{P} \to C$ is truthfully DSE-implemented by the game form $G = \langle N, \{A_i\}, g \rangle$ (with $A_i = \mathcal{P}$ for all $i \in N$), and for convenience let $N = \{1, \ldots, n\}$. Then for every $\succsim \in \mathcal{P}$ the action profile a^* in which $a_i^* = \succsim$ for all $i \in N$ is a dominant strategy equilibrium of the game $(G, \succsim)$ and $g(a^*) = f(\succsim)$. Suppose that a' is another dominant strategy equilibrium of $(G, \succsim)$. Then since both a_1^* and a_1' are dominant strategies for player 1 we have $g(a^*) \succsim_1 g(a_1', a_2^*, \ldots, a_n^*) \succsim_1 g(a^*)$; given the absence of indifference in the preference profiles it follows that $g(a^*) = g(a_1', a_2^*, \ldots, a_n^*)$. Similarly, since both a_2^* and a_2' are dominant strategies for player 2 we have $g(a_1', a_2^*, \ldots, a_n^*) \succsim_2 g(a_1', a_2', a_3^*, \ldots, a_n^*) \succsim_2 g(a_1', a_2^*, \ldots, a_n^*)$ and hence $g(a_1', a_2^*, \ldots, a_n^*) = g(a_1', a_2', a_3^*, \ldots, a_n^*)$. Continuing iteratively we deduce that $g(a^*) = g(a')$ and hence $g(a') = f(\succsim)$.

183.1 (*Example of non-DSE implementable rule*) Consider a preference profile $\succsim$ in which for some outcome a we have $x \succ_1 a \succ_1 a^*$ for all $x \notin \{a, a^*\}$, and for all $i \neq 1$ we have $a \succ_i x$ for all x. Let $\succsim_1'$ be a preference relation in which $a \succ_1' x \succ_1' a^*$ for all $x \notin \{a, a^*\}$. Now, using the revelation principle, in order for f to be DSE-implementable the preference profile $\succsim$ must be a dominant strategy equilibrium of the game $\langle G^*, \succsim \rangle$ defined in Lemma 181.4 *b*. But $f(\succsim) = a^*$ and $f(\succsim_{-1}, \succsim_i') = a$, so that $\succsim_1$ is not a dominant strategy for player 1 in $\langle G^*, \succsim \rangle$.

185.1 (*Groves mechanisms*)[1] We prove the claim in brackets at the end of the problem. If $x(\theta_{-j}, \overline{\theta}_j) = x(\theta_{-j}, \hat{\theta}_j)$ and $m_j(\theta_{-j}, \overline{\theta}_j) > m_j(\theta_{-j}, \hat{\theta}_j)$ then a player of type $\overline{\theta}_j$ is better off announcing $\hat{\theta}_j$ than $\overline{\theta}_j$. Thus if $x(\theta_{-j}, \overline{\theta}_j) = x(\theta_{-j}, \hat{\theta}_j)$ we must have $m_j(\theta_{-j}, \overline{\theta}_j) = m_j(\theta_{-j}, \hat{\theta}_j)$.

Now, denote $m_j^k = m_j(\theta_{-j}, \theta_j)$ for any value of θ_j such that $x(\theta_{-j}, \theta_j) = k$ ($\in \{0, 1\}$) and suppose that $x(\theta_{-j}, \theta_j) = 1$ and $x(\theta_{-j}, \theta_j') = 0$. Since it is a dominant strategy for player j with preference parameter $\theta_j'' = \gamma - \sum_{i \in N \setminus \{j\}} \theta_i$ to report θ_j'' he must be no better off if instead he reports θ_j' when the other players report θ_{-j}, so that $\theta_j'' - m_j^1 \geq -m_j^0$ or $\gamma - \sum_{i \in N \setminus \{j\}} \theta_i - m_j^1 \geq -m_j^0$. On the other hand, since, for any $\epsilon > 0$, it is a dominant strategy for player j with preference parameter $\theta_j'' = \gamma - \sum_{i \in N \setminus \{j\}} \theta_i - \epsilon$ to report θ_j'' he must be no better off if instead he reports θ_j when the other players report θ_{-j}, so that $-m_j^0 \geq \theta_j'' - m_j^1$ or $-m_j^0 \geq \gamma - \sum_{i \in N \setminus \{j\}} \theta_i - \epsilon - m_j^1$. Since this inequality holds for any $\epsilon > 0$ it follows that $-m_j^0 \geq \gamma - \sum_{i \in N \setminus \{j\}} \theta_i - m_j^1$. We conclude that $m_j^1 - m_j^0 = \gamma - \sum_{i \in N \setminus \{j\}} \theta_i$.

191.1 (*Implementation with two individuals*) The choice function is monotonic since $a \succsim_1 c$ and $c \succ_1' a$, and $b \succsim_2' c$ and $c \succ_2 b$.

Suppose that a game form G with outcome function g Nash-implements f. Then $(G, \succsim)$ has a Nash equilibrium, say (s_1, s_2), for which $g(s_1, s_2) = a$. Since (s_1, s_2) is a Nash equilibrium, $g(s_1, s_2') \precsim_2 a$ for all actions s_2' of player 2, so that $g(s_1, s_2') = a$ for all actions s_2' of player 2. That is, by choosing s_1, player 1 guarantees that the outcome is a. Since $a \succsim_1' b$, it follows that $(G, \succsim')$ has no Nash equilibrium (t_1, t_2) for which $g(t_1, t_2) = b$. We conclude that f is not Nash-implementable.

[1] *Correction to first printing of book*: "$x(\theta_{-j}, \theta_j') = 1$" on the last line of the problem should be "$x(\theta_{-j}, \theta_j') = 0$".

11 Extensive Games with Imperfect Information

203.2 (*Definition of* $X_i(h)$) Let $h = (a^1, \ldots, a^k)$ be a history, let $h^0 = \varnothing$, and let $h^r = (a^1, \ldots, a^r)$ for $1 \leq r \leq k-1$. Let $R(i)$ be the set of history lengths of subhistories of h after which player i moves; that is, let $R(i) = \{r\colon h^r \in I_i \text{ for some } I_i \in \mathcal{I}_i\}$ and denote by I_i^r the information set of player i that contains h^r when $r \in R(i)$. Then $X_i(h) = (I_i^{r_1}, a^{r_1+1}, \ldots, I_i^{r_\ell}, a^{r_\ell+1})$, where r_j is the jth smallest member of $R(i)$ and $\ell = |R(i)|$.

208.1 (*One-player games and principles of equivalence*)[1]

Inflation–deflation: The extensive game Γ is equivalent to the extensive game Γ' if Γ' differs from Γ only in that the player has an information set in Γ that is a union of information sets in Γ'. The additional condition in the general case (that any two histories in different members of the union have subhistories that are in the same information set of player i and player i's action at this information set is different in h and h') is always satisfied in a one-player game.

Coalescing of moves: Let h be a history in the information set I of the extensive game Γ, let $a \in A(h)$, and assume that (h, a) is not terminal. Let Γ' be the game that differs from Γ only in that the set of histories is changed so that for all $h' \in I$ the history (h', a) and the information set that contains (h', a) are deleted and every history of the type (h', a, b, h'') where $b \in A(h', a)$ is replaced by a history (h', ab, h'') where ab is a new action (that is not a member of $A(h')$), and the information sets and player's preferences are changed accordingly. Then Γ and Γ' are equivalent.

Now, by repeatedly applying inflation–deflation we obtain a game of perfect information. Repeated applications of the principle of coalescing of moves

[1] *Correction to first printing of book*: After "(but possibly with imperfect recall)" add "in which no information set contains both some history h and a subhistory of h".

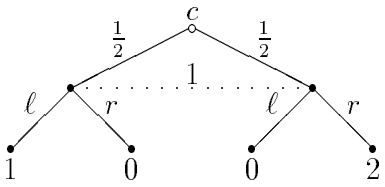

Figure 46.1 The one-player extensive game for the last part of Exercise 217.2.

yields a game with a single non-terminal history.

216.1 (*Example of mixed and behavioral strategies*) At the initial history choose A and B each with probability $\frac{1}{2}$; at the second information set choose ℓ.

217.1 (*Mixed and behavioral strategies and imperfect recall*) If player 1 uses the mixed strategy that assigns probability $\frac{1}{2}$ to $\ell\ell$ and probability $\frac{1}{2}$ to rr then she obtains the payoff of $\frac{1}{2}$ regardless of player 2's strategy. If she uses a behavioral strategy that assigns probability p to ℓ at the start of the game and probability q to ℓ at her second information set then she obtains the payoff $pqt + (1-p)(1-q)(1-t)$, where t is the probability with which player 2 chooses his left action. Thus by such a strategy she guarantees a payoff of only $\min\{pq, (1-p)(1-q)\}$, which is at most $\frac{1}{4}$ for any values of p and q.

217.2 (*Splitting information sets*) Suppose that the information set I^* of player 1 in the game Γ_2 is split into the two information sets I' and I'' in Γ_1. Let σ^* be a pure strategy Nash equilibrium of Γ_2 and define a profile σ' of pure strategies in Γ_1 by $\sigma'_i = \sigma^*_i$ for $i \neq 1$, $\sigma'_1(I') = \sigma'_1(I'') = \sigma^*(I^*)$, and $\sigma'_1(I) = \sigma^*_1(I)$ for every other information set I of player 1.

We claim that σ' is a Nash equilibrium of Γ_1. Clearly the strategy σ'_j of every player other than 1 is a best response to σ'_{-j} in Γ_1. As for player 1, any pure strategy in Γ_1 results in at most one of the information sets I' and I'' being reached, so that given σ'_{-1} any outcome that can be achieved by a pure strategy in Γ_1 can be achieved by a pure strategy in Γ_2; thus player 1's strategy σ'_1 is a best response to σ'_{-1}.

If Γ_2 contains moves of chance then the result does not hold: in the game in Figure 46.1 the unique Nash equilibrium is for the player to choose r. However, if the information set is split into two then the unique Nash equilibrium call for the player to choose ℓ if chance chooses the left action and r if chance chooses the right action.

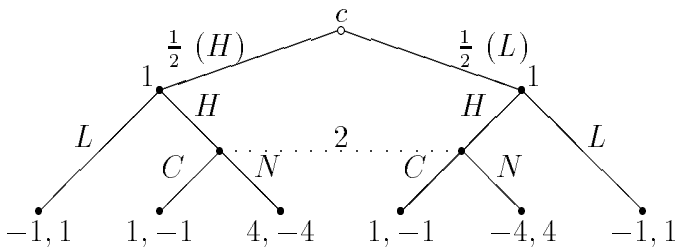

Figure 47.1 The extensive game for Exercise 217.3.

	C	N
LH	$0, \quad 0$	$-\frac{5}{2}, \quad \frac{5}{2}$
LL	$-1, \quad 1$	$-1, \quad 1$
HL	$0, \quad 0$	$\frac{3}{2}, -\frac{3}{2}$
HH	$1, -1$	$0, \quad 0$

Figure 47.2 The strategic form of the extensive game in Figure 47.1.

217.3 (*Parlor game*) This (zerosum) extensive game is shown in Figure 47.1. The strategic form of this game is given in Figure 47.2. First note that the strategies LH and LL are both strictly dominated by HH. (I.e. if player 1 gets the high card she is better off not conceding.) Now, there is a unique Nash equilibrium, in which the mixed strategy of player 1 assigns probability $\frac{2}{5}$ to HL and probability $\frac{3}{5}$ to HH and player 2 concedes with probability $\frac{3}{5}$. (In behavioral strategies this equilibrium is: player 1 chooses H when her card is H and chooses H with probability $\frac{3}{5}$ and L with probability $\frac{2}{5}$ when her card is L; player 2 concedes with probability $\frac{3}{5}$.)

12 Sequential Equilibrium

226.1 (*Example of sequential equilibria*) Denote player 1's strategy by (α, β, ζ). In all sequential equilibria:

- If $\beta > \zeta$ then player 2 chooses L and hence $\beta = 1$; (M, L) is indeed a sequential equilibrium strategy profile.
- If $\beta < \zeta$ then player 2 chooses R, so that player 1 chooses L and $\beta = \zeta = 0$, a contradiction.
- If $\beta = \zeta > 0$ then player 2 must choose L with probability $\frac{1}{2}$, in which case player 1 is better off choosing L, a contradiction.
- If $\beta = \zeta = 0$ then player 2's strategy $(\delta, 1-\delta)$ has to be such that $1 \geq 3\delta - 2(1-\delta) = 5\delta - 2$ or $\frac{3}{5} \geq \delta$, and $1 \geq 2\delta - (1-\delta) = 3\delta - 1$ or $\frac{2}{3} \geq \delta$. For each $0 < \delta \leq \frac{3}{5}$ the strategy is supported by the belief $(\frac{1}{2}, \frac{1}{2})$ of player 2. For $\delta = 0$ the strategy is supported by any belief $(p, 1-p)$ with $p \leq \frac{1}{2}$.

In summary, there are two types of sequential equilibria: one in which the strategy profile is $((0,1,0),(1,0))$ and player 2's belief is $(1,0)$, and one in which the strategy profile is $((1,0,0),(\delta,1-\delta))$ for some $\delta \in [0, \frac{3}{5}]$ and player 2's belief is $(\frac{1}{2}, \frac{1}{2})$ for $\delta > 0$ and $(p, 1-p)$ for some $p \leq \frac{1}{2}$ for $\delta = 0$.

227.1 (*One deviation property for sequential equilibrium*) (This proof is taken from Hendon, Jacobsen, and Sloth (1993).)

First note that by the assumption of perfect recall, if the information set I_i' of player i contains a history $(h, a^1, \ldots, a^k)$ for which $h \in I_i$ then all histories in I_i' are of the form $(h', b^1, \ldots, b^m)$ for some $h' \in I_i$, where the sequence of actions of player i in the sequence $(a^1, \ldots, a^k)$ is the same as the sequence of actions of player i in the sequence $(b^1, \ldots, b^m)$

Now suppose that (β, μ) is a consistent assessment, let β_i' be a strategy of player i, let $\beta' = (\beta_{-i}, \beta_i')$, let I_i and I_i' be information sets of player i, and let $h = (\hat{h}, a', a'')$ be a terminal history, where a' and a'' are sequences of actions, $\hat{h} \in I_i$, and $(\hat{h}, a') \in I_i'$. We begin by showing that $O(\beta', \mu|I_i)(h) = O(\beta', \mu|I_i')(h) \cdot \Pr(\beta', \mu|I_i)(I_i')$. If $\Pr(\beta', \mu|I_i)(I_i') = 0$ then this equality certainly holds, so suppose that $\Pr(\beta', \mu|I_i)(I_i') > 0$. Then we have

$$O(\beta', \mu|I_i)(h) = \mu(I_i)(\hat{h}) \cdot P_{\beta'}(a', a'')$$

and

$$O(\beta', \mu|I_i')(h) = \mu(I_i')(\hat{h}, a') \cdot P_{\beta'}(a''),$$

where $P_{\beta'}(a)$ is the product of the probabilities assigned by β' to the sequence a of actions. Now for all $h' \in I_i'$ let $\overline{h}(h')$ be the subhistory of h' in I_i (existence and uniqueness follows from perfect recall). Let $h' \setminus \overline{h}(h')$ be the part of h' subsequent to I_i. Then,

$$\Pr(\beta', \mu|I_i)(I_i') = \sum_{h' \in I_i'} \mu(I_i)(\overline{h}(h')) \cdot P_{\beta'}(h' \setminus \overline{h}(h')).$$

Since (β, μ) is consistent there is a sequence of completely mixed assessments (β^n, μ^n) with $\mu^n \to \mu$ and $\beta^n \to \beta$ as $n \to \infty$ and for all n the belief μ^n is derived from β^n using Bayes' rule. For each n we have

$$\mu^n(I_i')(\hat{h}, a') = \frac{\mu^n(I_i)(\hat{h}) \cdot P_{\beta'}(a')}{\sum_{h' \in I_i} \mu^n(I_i)(\overline{h}(h')) \cdot P_{\beta'}(h' \setminus \overline{h}(h'))}$$

since $\Pr(\beta', \mu|I_i)(I_i') > 0$. Taking the limit as $n \to \infty$ and using $P_{\beta'}(a', a'') = P_{\beta'}(a') \cdot P_{\beta'}(a'')$ we conclude that $O(\beta', \mu|I_i)(h) = O(\beta', \mu|I_i')(h) \cdot \Pr(\beta', \mu|I_i)(I_i')$.

To show the one deviation property, use backwards induction. Suppose that (β, μ) is a consistent assessment with the property that no player has an information set at which a change in his action (holding the remainder of his strategy fixed) increases his expected payoff conditional on reaching that information set. Take an information set I_i of player i and suppose that β_i is optimal conditional on reaching any of the information sets I_i' of player i that immediately follow I_i. We need to show that β_i is optimal conditional on reaching I_i. Suppose that player i uses the strategy β_i'. Let $\beta' = (\beta_{-i}, \beta_i')$, let $\mathcal{F}(I_i)$ be the set of information sets of player i that immediately follow I_i, and let $Z(I_i)$ be the set of terminal histories that have subhistories in I_i. Then player i's expected payoff conditional on reaching I_i is the sum of his payoffs to histories that do not reach another of his information sets, say E_i, and

$$\sum_{I_i' \in \mathcal{F}(I_i)} \sum_{h \in Z(I_i')} O(\beta', \mu|I_i)(h) \cdot u_i(h).$$

This is equal, using the equality in the first part of the problem, to

$$E_i + \sum_{I_i' \in \mathcal{F}(I_i)} \sum_{h \in Z(I_i')} O(\beta', \mu | I_i')(h) \cdot \Pr(\beta', \mu | I_i)(I_i') \cdot u_i(h),$$

which is equal to

$$E_i + \sum_{I_i' \in \mathcal{F}(I_i)} \Pr(\beta', \mu | I_i)(I_i') \cdot \mathrm{E}_{(\beta', \mu)}[u_i | I_i'],$$

where $\mathrm{E}_{(\beta', \mu)}[u_i | I_i']$ is the expected payoff under (β', μ) conditional on reaching I_i', which by the induction assumption is at most

$$E_i + \sum_{I_i' \in \mathcal{F}(I_i)} \Pr(\beta', \mu | I_i)(I_i') \cdot \mathrm{E}_{(\beta, \mu)}[u_i | I_i'].$$

Now, again using the equality in the first part of the problem, this is equal to

$$\mathrm{E}_{((\beta_{-i}, \hat{\beta}_i), \mu)}[u_i | I_i],$$

where $\hat{\beta}_i$ is the strategy of player i in which player i uses β_i' at I_i and β_i elsewhere. Thus β_i is optimal conditional on reaching I_i.

229.1 (*Non-ordered information sets*) The three sequential equilibria are:

- *Strategies* $\beta_1(s) = 1$, $\beta_2(d) = 1$, $\beta_3(s) = 1$.
  Beliefs $\mu_1(a) = 1$, $\mu_2(a, c) = \mu_2(b, e) = \frac{1}{2}$, $\mu_3(b) = 1$.
- *Strategies* $\beta_1(c) = 1$, $\beta_2(\ell) = 1$, $\beta_3(e) = 1$.
  Beliefs $\mu_1(a) = 1$, $\mu_2(a, c) = \mu_2(b, e) = \frac{1}{2}$, $\mu_3(b) = 1$.
- *Strategies* $\beta_1(c) = 1$, $\beta_2(r) = 1$, $\beta_3(e) = 1$.
  Beliefs $\mu_1(a) = 1$, $\mu_2(a, c) = \mu_2(b, e) = \frac{1}{2}$, $\mu_3(b) = 1$.

It is straightforward to check that each of these assessments satisfies sequential rationality and consistency.

The first equilibrium has the following undesirable feature. Player 2's strategy d is optimal only if he believes that each of the two histories in his information set occurs with probability $\frac{1}{2}$. If he derives such a belief from beliefs about the behavior of players 1 and 3 then he must believe that player 1 chooses c with positive probability and player 3 chooses e with positive probability. But then it is no longer optimal for him to choose d: ℓ and r both yield him 2, while

d yields less than 2. That is, any alternative strategy profile that rationalizes player 2's belief in the sense of structural consistency makes player 2's action in the sequential equilibrium suboptimal.

Nevertheless, player 2's strategy can be rationalized by another explanation of the reason for reaching the information set. Assume that player 2 believes that players 1 and 3 attempted to adhere to their behavioral strategies but made errors in carrying out these strategies. Then the fact that he believes that there is an equal probability that each of them made a mistake does not mean that he has to assign a positive probability to a mistake in the future.

234.2 (*Sequential equilibrium and PBE*) Since (β, μ) is a sequential equilibrium there is a sequence $(\beta^n, \mu^n)_{n=1}^{\infty}$ of assessments that converges to (β, μ) and has the properties that each strategy profile β^n is completely mixed and each belief system μ^n is derived from β^n using Bayes' law. For each $h \in H$, $i \in P(h)$, and $\theta_i \in \Theta_i$ let $\sigma_i^n(\theta_i)(h) = \beta_i^n(I(\theta_i, h))$ for each value of n. Given these (completely mixed) strategies define a profile (μ_i^n) of beliefs in the Bayesian extensive game that satisfies the last three conditions in Definition 232.1. It is straightforward to show that $\mu^n(I(\theta_i, h))(\theta, h) = \Pi_{j \in N \setminus \{i\}} \mu_j^n(h)(\theta_j)$ for each value of n. This equality and the properties of (μ_i^n) are preserved in the limit, so that $\mu(I(\theta_i, h))(\theta, h) = \Pi_{j \in N \setminus \{i\}} \mu_j(h)(\theta_j)$. Thus by the sequential rationality of the sequential equilibrium, $((\sigma_i), (\mu_i))$ is sequentially rational and hence a perfect Bayesian equilibrium.

237.1 (*Bargaining under imperfect information*) Refer to the type of player 1 whose valuation is v as *type* v. It is straightforward to check that the following assessment is a sequential equilibrium: type 0 always offers the price of 2 and type 3 always offers the price of 5. In both periods player 2 accepts any price at most equal to 2 and rejects all other prices (regardless of the history). If player 2 observes a price different from 5 in either period then he believes that he certainly faces type 0. (Thus having rejected a price of 5 in the first period, which he believed certainly came from type 3, he concludes, in the event that he observes a price different from 5 in the second period, that he certainly faces type 0.)

Comment There are other sequential equilibria, in which both types offer a price between 3 and 3.5, which player 2 immediately accepts.

238.1 (*PBE is SE in Spence's model*) It is necessary to show only that the assessments are consistent. Consider the pooling equilibrium. Suppose that

a type θ_1^L worker chooses e^* with probability $1-\epsilon$ and distributes the remaining probability ϵ over other actions, while a type θ_1^H worker chooses e^* with probability $1-\epsilon^2$ and distributes the remaining probability ϵ^2 over other actions. The employer's belief that these completely mixed strategies induce converges to the one in the equilibrium as $\epsilon \to 0$, so that the equilibrium assessment is indeed consistent. A similar argument shows that the separating equilibrium is a sequential equilibrium.

243.1 (*PBE of chain-store game*) The challengers' beliefs are initially correct and action-determined, and it is shown in the text that the challengers' strategies are sequentially rational, so that it remains to show that the chain-store's strategy is sequentially rational and that the challengers' beliefs satisfy the condition of Bayesian updating.

Sequential rationality of regular chain-store's strategy:

- If $t(h) = K$ then the regular chain-store chooses C, which is optimal.
- Suppose that $t(h) = k \leq K-1$ and $\mu_{CS}(h)(T) \geq b^{K-k}$. Then if the chain-store chooses C it obtains 0 in the future. If it chooses F then challenger $k+1$ believes that the probability that the chain-store is tough is $\max\{b^{K-k}, \mu_{CS}(h)(T)\}$ and stays out. Thus if the chain-store chooses F then it obtains -1 against challenger k and a against challenger $k+1$. Thus it is optimal to choose F.
- Suppose that $t(h) = k \leq K-1$ and $\mu_{CS}(h)(T) < b^{K-k}$. Then if the chain-store chooses C it obtains 0 in the future. If it chooses F then challenger $k+1$ believes that the probability that the chain-store is tough is $\max\{b^{K-k}, \mu_{CS}(h)(T)\} = b^{K-k}$ and chooses *Out* with probability $1/a$. Thus if the chain-store chooses F against challenger k and challenger $k+1$ chooses *Out* then the chain-store obtains a total payoff of $-1 + a \cdot (1/a) = 0$ when facing these two challengers. If the chain-store chooses F against challenger k and challenger $k+1$ chooses *In* then the chain-store randomizes in such a way that it obtains an expected payoff of 0 regardless of its future actions. Thus the chain-store's expected payoff if it chooses F against challenger k is zero, so that it is optimal for it to randomize between F and C.

Sequential rationality of tough chain-store's strategy: If the tough chain-store chooses C after any history then all future challengers enter. Thus it is optimal for the tough chain-store to choose F.

Bayesian updating of beliefs:

- If $k \leq K-1$ and $\mu_{CS}(h)(T) \geq b^{K-k}$ then both types of chain-store fight challenger k if it enters. Thus the probability $\mu_{CS}(h, h_k)(T)$ assigned by challenger $k+1$ is $\mu_{CS}(h)(T)$ when $h_k = (In, F)$.

- If $k \leq K-1$ and $\mu_{CS}(h)(T) < b^{K-k}$ then the tough chain-store fights challenger k if it enters and the regular chain-store accommodates with positive probability $p_k = (1 - b^{K-k})\mu_{CS}(h)(T)/((1 - \mu_{CS}(h)(T))b^{K-k})$. Thus in this case

$$\mu_{CS}(h, h_k)(T) = \frac{\mu_{CS}(h)(T)}{\mu_{CS}(h)(T) + (1 - \mu_{CS}(h)(T))p_k} = b^{K-k}$$

  if $h_k = (In, F)$.

- If $\mu_{CS}(h)(T) = 0$ or $h_k = (In, C)$, $k \leq K-1$, and $\mu_{CS}(h)(T) < b^{K-k}$ then we have $\mu_{CS}(h, h_k)(T) = 0$ since only the regular chain-store accommodates in this case.

- If $h_k = (In, C)$, $k \leq K-1$, and $\mu_{CS}(h)(T) \geq b^{K-k}$ then neither type of chain-store accommodates entry, so that if C is observed challenger $k+1$ can adopt whatever belief it wishes; in particular it can set $\mu_{CS}(h, h_k)(T) = 0$.

246.2 (*Pre-trial negotiation*) The signaling game is the Bayesian extensive game with observable actions $\langle \Gamma, (\Theta_i), (p_i), (u_i) \rangle$ in which Γ is a two-player game form in which player 1 first chooses either 3 or 5 and then player 2 chooses either *Accept* or *Reject*; $\Theta_1 = \{Negligent, Not\}$, Θ_2 is a singleton, and $u_i(\theta, h)$ takes the values described in the problem.

The game has no sequential equilibrium in which the types of player 1 make different offers. To see this, suppose that the negligent type offers 3 and the non-negligent type offers 5. Then the offer of 3 is rejected and the offer of 5 is accepted, so the negligent player 1 would be better off if she offered 5. Now suppose that the negligent type offers 5 and the non-negligent type offers 3. Then both offers are accepted and the negligent type would be better off if she offered 3.

The only sequential equilibria in which the two types of player 1 make the same offer are as follows.

- If $p_1(Not) \geq \frac{2}{5}$ then the following assessment is a sequential equilibrium. Both types of player 1 offer the compensation of 3 and player 2 accepts any offer. If the compensation of 3 is offered then player 2 believes that player 1 is not negligent with probability $p_1(Not)$; if the compensation 5 is offered then player 2 may hold any belief about player 1. (The condition $p_1(Not) \geq \frac{2}{5}$ is required in order for it to be optimal for player 2 to accept when offered the compensation 3.)

- For any value of $p_1(Not)$ the following assessment is a sequential equilibrium. Both types of player 1 offer the compensation 5; player 2 accepts an offer of 5 and rejects an offer of 3. If player 2 observes the offer 3 then he believes that player 1 is not negligent with probability at most $\frac{2}{5}$.

Consider the case in which $p_1(Not) > \frac{2}{5}$. The second type of equilibrium involves the possibility that if player 1 offers only 3 then the probability assigned by player 2 to her being negligent is increasing. A general principle that excludes such a possibility emerges from the assumption that whenever it is optimal for a negligent player 1 to offer the compensation 3 it is also optimal for a non-negligent player 1 to do so. Thus if the out-of-equilibrium offer 3 is observed a reasonable restriction on the belief is that the relative probability of player 1 being non-negligent should increase and thus exceed $\frac{2}{5}$. However, if player 2 holds such a belief then his planned rejection is no longer optimal.

252.2 (*Trembling hand perfection and coalescing of moves*) In the original game the history (L, R) is an outcome of a trembling hand perfect equilibrium in which player 1 chooses (L, r) and player 2 chooses R. If we coalesce player 1's moves then we get the game in which player 1 chooses between the three actions L, $R\ell$, and Rr. In this game the only trembling hand perfect equilibrium is (Rr, R).

Comment If the game is modifies so that the payoffs of player 2 to the history (L, R) and (R, r) remain positive but are different then coalescing player 1's moves affects the players' equilibrium payoffs.

253.1 (*Example of trembling hand perfection*) The extensive form of the game is given in Figure 56.1.

The reduced strategic form is shown in Figure 56.2. The only strategies that are not weakly dominated are $1g$ for player 1 and $2g$ for player 2. Thus by Proposition 248.2 the strategy profile $(1g, 2g)$ is the unique trembling hand perfect equilibrium of the strategic form of the game.

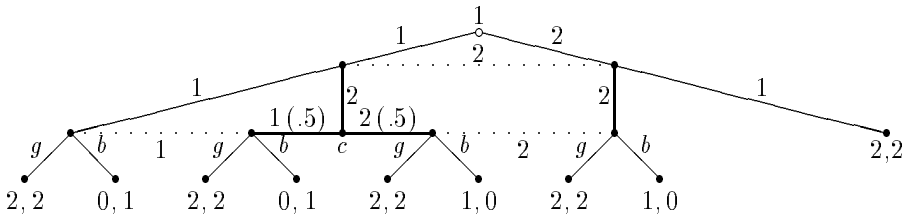

Figure 56.1 The extensive form of the game in Exercise 253.1

	1	$2g$	$2b$
$1g$	$2, 2$	$2, 2$	$\frac{3}{2}, 1$
$1b$	$0, 1$	$1, \frac{3}{2}$	$\frac{1}{2}, 1\%over2$
2	$2, 2$	$2, 2$	$1, 0$

Figure 56.2 The reduced strategic form of the game in Figure 56.1.

We now argue that $((1, g), (2, g))$ is not a trembling hand perfect equilibrium of the extensive game. By definition, a trembling hand perfect equilibrium of the extensive game corresponds to a trembling hand perfect equilibrium of the agent strategic form of the game. Consider a completely mixed strategy profile of the agent strategic form of the game. Assume that the probability with which player 1's second agent chooses b is at least as large as the probability with which player 2's second agent chooses b. Then the only best response of player 1's first agent is to choose 2. To see this, let p_i be the probability with which player i's first agent chooses i and let q_i be the probability that player i's second agent chooses g. Then player 1's payoff if her first agent chooses 1 is

$$(1 - p_2) \cdot 2q_1 + p_2 \cdot [\tfrac{1}{2} \cdot 2q_1 + \tfrac{1}{2}(2q_2 + 1 - q_2)]$$

while her payoff if her first agent chooses 2 is

$$(1 - p_2) \cdot 2 + p_2 \cdot [2q_2 + 1 - q_2].$$

The difference between the first and second of these payoffs is

$$2(1 - p_2)(q_1 - 1) + p_2 \cdot [q_1 - q_2 - \tfrac{1}{2}(1 - q_2)] < 0$$

if $q_2 \geq q_1$. A symmetric argument applies to player 2's first agent. Thus given any completely mixed strategy profile, for at least one player it is better for that players's first agent to choose the other player.

Interpretation: Trembling hand perfect equilibrium in the strategic form captures the idea that each player is concerned about (small) mistakes that his opponent may make, which leads each player in this game to choose himself to be the one to make the decision. Trembling hand perfect equilibrium in the extensive game allows for the fact that the player may make mistakes himself in carrying out his strategy later in the game, which in this game, given that errors by oneself are more costly than errors by one's opponent, militates against choosing oneself to be the decision-maker.

13 The Core

259.3 (*Core of production economy*) First suppose that the payoff profile x is a member of the set given. If S does not contain the capitalist then $v(S) = 0$, so certainly $x(S) \geq v(S)$. If S does contain the capitalist then $x(S) = f(w) - \sum_{i \in N \setminus S} x_i \geq f(w) - (w + 1 - |S|)(f(w) - f(w-1))$, which is at least $f(|S|)$ by the concavity of f. Thus x is in the core.

Now suppose that x is a feasible payoff profile for which $x_i > f(w) - f(w-1)$ for some $i \neq c$. Then $x(N \setminus \{i\}) = f(w) - x_i < f(w) - (f(w) - f(w-1)) = f(w-1) = v(N \setminus \{i\})$, so that x is not in the core.

In each payoff profile in the core each worker receives not more than his marginal product when all workers are employed, and the capitalist receives the residual.

260.2 (*Market for indivisible good*) Let x be a payoff profile in the core, let b be a buyer whose payoff is minimal among the payoffs of all the buyers, and let ℓ be a seller whose payoff is minimal among the payoffs of all the sellers. Then $x_b + x_\ell \geq v(\{b, \ell\}) = 1$; since $|L| = v(N) \geq |B| x_b + |L| x_\ell = |L|(x_b + x_\ell)$ it follows that $x_b + x_\ell = 1$ and $x_i = x_j$ if i and j are both buyers or if they are both sellers. Thus the core is the set of all payoff profiles in which for some $\alpha \in [0, 1]$ every buyer receives the payoff α and every seller receives the payoff $1 - \alpha$. That is, any split of the surplus is possible in this case; the only impact of the competition between buyers and between sellers is that all pairs must split in the same way.

260.4 (*Convex games*) Let $S^* = \{i_1, \ldots, i_{|S^*|}\}$ be any coalition, with $i_1 < \cdots < i_{|S^*|}$. Then $x_{i_1} = v(S_{i_1} \cup \{i_1\}) - v(S_{i_1}) \geq v(\{i_1\})$ (take $S = S_{i_1}$ and $T = \{i_1\}$ in the definition of convexity). But then $x_{i_1} + x_{i_2} \geq v(\{i_1\}) + v(S_{i_2} \cup \{i_2\}) - v(S_{i_2}) \geq v(\{i_1, i_2\})$ (take $S = S_{i_2}$ and $T = \{i_1, i_2\}$ in the definition of convexity).

Continuing similarly we reach the conclusion that $x_{i_1} + \ldots + x_{i_{|S^*|}} \geq v(S^*)$. Further, $\sum_{i \in N} x_i = v(N)$, so that x is in the core of $\langle N, v \rangle$.

261.1 (*Simple games*)

a. For each $i \in N$ let S_i be a winning coalition that does not contain i; let x be a payoff profile in the core. Then

$$x(N \setminus \{i\}) \geq x(S_i) \geq v(S_i) = 1,$$

so that $\sum_{i \in N} x(N \setminus \{i\}) \geq |N|$. On the other hand

$$\sum_{i \in N} x(N \setminus \{i\}) = (|N| - 1) \sum_{i \in N} x_i = |N| - 1,$$

a contradiction.

b. Let V be the set of veto players. Let x be a nonnegative feasible payoff profile for which $x_i = 0$ for all $i \in N \setminus V$. If S is not a winning coalition then $v(S) = 0$ so that certainly $x(S) \geq v(S)$; if S is a winning coalition then $x(S) = 1 = v(S)$. Thus x is in the core. Now, if x is in the core then since $v(S) \geq 0$ for all S we have $x_i \geq 0$ for all $i \in N$. Let x be a feasible payoff profile for which $x_i > 0$ for some $i \in N \setminus V$. Let S be a winning coalition that does not include i. Then $x(S) < 1 = v(S)$, so that x is not in the core.

261.2 (*Zerosum games*) If $\langle N, v \rangle$ is zerosum and x is in the core of $\langle N, v \rangle$ then for any coalition S we have $x(S) \geq v(S)$ and $x(N \setminus S) \geq v(N \setminus S)$; since $x(S) + x(N \setminus S) = x(N) = v(N) = v(S) + v(N \setminus S)$ it follows that $x(S) = v(S)$. Thus for all disjoint coalitions S and T we have $v(S) + v(T) = x(S) + x(T) = x(S \cup T) = v(S \cup T)$. Hence $\langle N, v \rangle$ is additive.

261.3 (*Pollute the lake*)

a. Let S be a coalition and let $|S| = s$. The payoff of S is minimized if none of the members of $N \setminus S$ treats its waste. In this case the payoff of S if k of its members treat their waste is $-s(n-k)c - kb$. Thus if $sc \geq b$ then the payoff of S is maximized when all members of S treat their waste, yielding S a payoff of $-s(n-s)c - sb$, and if $sc \leq b$ then the payoff of S is maximized when no member of S treats its waste, yielding S a payoff of $-snc$. Thus

$$v(S) = \begin{cases} -snc & \text{if } s < b/c \\ -s[(n-s)c + b] & \text{if } s \geq b/c. \end{cases}$$

b. First we argue that since the game is symmetric the core is nonempty if and only if it contains the payoff profile $x = (-b, \ldots, -b)$. To see this, suppose that x is not in the core. Then for some integer k such that $v(S) > -kb$ for every coalition S with $|S| = k$. Now let $y \neq x$ be a feasible payoff profile. Then there exists some coalition T with $|T| = k$ and $y(T) < -kb = v(T)$. Thus y is not in the core.

Now, if $|S| = s \leq b/c$ then $x(S) = -sb \geq -snc = v(S)$ and if $|S| = s > b/c$ then $x(S) = -sb \geq -s[(n-s)c + b] = v(S)$ and $x(N) = -nb = v(N)$ (by the assumption that $b \leq nc$). Thus x is in the core of the game, which consequently is always nonempty.

The core is a singleton if and only if $b = nc$. To show this, first suppose that $b = nc$ and $x \neq (-b, \ldots, -b)$. Then $x_i < -b$ for some $i \in N$, so that $x(\{i\}) < v(\{i\}) = -nc = -b$ (since $c \leq b$); thus x is not in the core. Conversely, if $b < nc$ and $x = (-b, \ldots, -b)$ then $x(S) > v(S)$ whenever $|S| < n$, so that the core contains payoff profiles different from x.

c. Under the assumptions in the exercise a coalition is pessimistic about the outcome when it deviates, and consequently does so only when it is *sure* that it can increase its payoff from doing so. The value of $v(S)$ for each $S \neq N$ is smaller than it is under alternative assumptions, causing the core to be larger than it is under alternative assumptions.

263.2 (*Game with empty core*) Let $\lambda_{\{1,2\}} = \lambda_{\{1,3\}} = \lambda_{\{1,4\}} = \frac{1}{3}$ and $\lambda_{\{2,3,4\}} = \frac{2}{3}$. Then (λ_S) is a balanced collection of weights; since $\frac{1}{3}v(\{1,2\}) + \frac{1}{3}v(\{1,3\}) + \frac{1}{3}v(\{1,4\}) + \frac{2}{3}v(\{2,3,4\}) = \frac{5}{4} > v(N)$ the game is not balanced and thus (by the Bondareva–Shapley theorem) has an empty core.

265.2 (*Syndication in a market*)

a. We have $v(S) = \min\{2|S \cap \{1,2\}|, |S \cap \{3,4,5\}|\}$ for each coalition S. If x is in the core then $x_1 + x_i + x_j \geq 2$ whenever $\{i,j\} \subseteq \{3,4,5\}$, so that $3x_1 + 2(x_3 + x_4 + x_5) \geq 6$ and hence $x_1 \geq 2x_2$ (using $x_3 + x_4 + x_5 = 3 - x_1 - x_2$). Similarly $x_2 \geq 2x_1$, so that $x_1 = x_2 = 0$. We also require $x_1 + x_i \geq 1$ if $i \in \{3,4,5\}$, so that the core is $\{(0,0,1,1,1)\}$.

b. Let the players be 1, 2, and s (the syndicate). We have $v(\{1,s\}) = v(\{2,s\}) = 2$, $v(N) = 3$, and $v(S) = 0$ otherwise. The core is the set of feasible payoff profiles for which $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 1$. Thus the core predicts that the members of the syndicate are never better off, and may be worse off. An interpretation is that the fact that 3, 4, and 5 always act as a block dulls the competition between 1 and 2, who cannot now compete with

each other by forming (efficient) coalitions consisting of only two of the three members of 3, 4, and 5. (The payoff profile $(1, 1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is not in the core of the unsyndicated market since the coalition $\{1, 3, 4\}$ can obtain 2 units of payoff.)

267.2 (*Existence of competitive equilibrium in market*) First note that the two sets are nonempty and convex and their interiors are disjoint, so that indeed they can be separated. Thus there exists $(\alpha, \beta) \in \mathbb{R}^\ell \times \mathbb{R}$, not equal to 0, such that

$$\alpha \cdot z + \beta y \leq \alpha \cdot \sum_{i \in N} z_i^* + \beta \sum_{i \in N} f_i(z_i^*) \text{ for all } (z, y) \in X.$$

Since $(\sum_{i \in N} z_i^* + 1_j, \sum_{i \in N} f_i(z_i^*)) \in X$, where 1_j is the jth unit vector, we have $\alpha_j \leq 0$ for all j. We now show that $\beta > 0$. Since $\sum_{i \in N} \omega_i > 0$ there exists $\epsilon \in \mathbb{R}^\ell_{++}$ and $\delta > 0$ such that $(\sum_{i \in N} z_i^* - \epsilon, \sum_{i \in N} f_i(z_i^*) - \delta) \in X$, so that $-\alpha \cdot \epsilon - \beta\delta \leq 0$ or $\beta\delta \geq -\alpha \cdot \epsilon$. If $\alpha = 0$ then we conclude that $\beta \geq 0$, and since $(\alpha, \beta) \neq 0$ it follows that $\beta > 0$. If $\alpha_j < 0$ for some j then we conclude directly that $\beta > 0$.

Now let $p = -\alpha/\beta \geq 0$. Since $(\sum_{i \in N} z_i^* - z_k^* + z_k, \sum_{i \in N} f_i(z_i^*) - f_k(z_k^*) + f_k(z_k)) \in X$ for any $z_k \in \mathbb{R}^\ell_+$ we have

$$f_k(z_k^*) - pz_k^* \geq f_k(z_k) - pz_k \text{ for all } z_k \in \mathbb{R}^\ell_+,$$

so that $(p, (z_i^*)_{i \in N})$ is a competitive equilibrium.

Comment This is not an exercise in game theory.

268.1 (*Core convergence in production economy*) In all elements of the core the payoff of every player $i \neq 1$ is at most $f(1, k) - f(1, k-1)$ (see Exercise 259.3). Now, the concavity of $f(1, k)$ implies that $k(f(1, k) - f(1, k-1)) \leq 2(f(1, k) - f(1, k/2))$ (since

$$\begin{aligned} f(1, k) - f(1, k/2) &= \sum_{j=k/2+1}^{k} (f(1, j) - f(1, j-1)) \\ &\geq \sum_{j=k/2+1}^{k} (f(1, k) - f(1, k-1)) \\ &\geq (k/2)[f(1, k) - k(1, k-1)]). \end{aligned}$$

Since f is bounded we have $f(1, k) - f(1, k/2) \to 0$, establishing the result.

Interpretation: Competition between the workers drives their payoff down to their marginal product, which declines to zero, so that the single capitalist gets all the surplus.

274.1 (*Core and equilibria of exchange economy*) We first claim that the only competitive price is $(p_1, p_2) = (\frac{1}{2}, \frac{1}{2})$. To see this, suppose that $p_1 > p_2$; then each agent of type 1 demands none of good 1 and each agent of type 2 demands less than $\frac{1}{2}$ a unit of good 1, so that the aggregate demand for good 1 is less than the supply. If $p_1 < p_2$ then each agent of type 1 demands 1 unit of good 1 and each agent of type 2 demands more than $\frac{1}{2}$ a unit of good 1, so that the aggregate demand for good 1 exceeds the supply. An allocation is competitive if each agent i of type 1 obtains the bundle $(y_i, 1 - y_i)$ for some $y_i \in [0, 1]$ and each agent of type 2 obtains the bundle $(\frac{1}{2}, \frac{1}{2})$, where $\sum_{i \text{ of type } 1} y_i = k/2$.

Now consider the core. First suppose that $k = 1$. In order for the allocation $((s, t), (1 - s, 1 - t))$ to be in the core we require $s + t \geq 1$ (considering the coalition $\{1\}$) and $1 - s = 1 - t$ (considering the coalition $\{1, 2\}$). Thus the core consists of all allocations $((s, s), (1 - s, 1 - s))$ for which $s \geq \frac{1}{2}$.

Now suppose that $k \geq 2$. We claim that the core of kE is the set of competitive allocations. We show this as follows. Let x be an allocation in the core.

Step 1. For each agent i of type 2 we have $x_i = (y_i, y_i)$ for some $y_i \in [0, 1]$. The argument is straightforward.

Step 2. Each agent obtains the same payoff. The argument is the same as that for Lemma 272.2 (the equal treatment result).

Step 3. Each agent of type 2 obtains the same bundle. This follows from Steps 1 and 2.

Step 4. Each agent of type 2 obtains the bundle $(\frac{1}{2}, \frac{1}{2})$. By Steps 1, 2, and 3 each agent of type 2 obtains the same bundle (y, y) with $y \leq \frac{1}{2}$. Suppose that $y < \frac{1}{2}$. Then each agent of type 1 obtains the payoff $2(1 - y)$. Consider a coalition S that consists of one agent of type 1 and two agents of type 2. The endowment of S is $(1, 2)$, so that it is feasible to give the agent of type 1 the bundle $(1 - 2y - 2\epsilon, 2 - 2y - 2\epsilon)$ and each agent of type 2 the bundle $(y + \epsilon, y + \epsilon)$ if $\epsilon > 0$ is small enough. In this allocation the payoff of each agent exceeds his payoff in the original allocation if ϵ is small enough, establishing the result.

Finally, it is easy to show that each allocation in which each agent i of type 1 obtains the bundle $(y_i, 1 - y_i)$ for some $y_i \in [0, 1]$ and each agent of type 2 obtains the bundle $(\frac{1}{2}, \frac{1}{2})$ is indeed in the core.

14 Stable Sets, the Bargaining Set, and the Shapley Value

280.1 (*Stable sets of simple games*) Let Y be the set of imputations described in the problem. To show internal stability let $y \in Y$ and suppose that $z \succ_S y$ for some $z \in Y$. Then $z_i > y_i \geq 0$ for all $i \in S$, so that $z(S) > y(S)$. Since $z \in Y$ we have $S \subseteq T$; since S is winning and T is minimal winning we have $T \subseteq S$. Thus $z(S) = y(S)$, a contradiction. To show external stability let $z \in X \setminus Y$. Then $\sum_{i \in T} z_i < 1$ so that there exists $y \in Y$ such that $y \succ_T z$.

280.2 (*Stable set of market for indivisible good*)

Internal stability: Let $y \in Y$ and suppose that $z \in Y$ with $z_i > y_i$ for all $i \in S$. Then $S \subseteq L$ or $S \subseteq B$; but then $v(S) = 0$.

External stability: Let $z \in X \setminus Y$. Let i be the buyer whose payoff is lowest among all buyers, let j be the seller whose payoff is lowest among all sellers, let z^b be the average payoff of the buyers in z, and let z^ℓ be the average payoff of the sellers. Since $|B|z^b + |L|z^\ell = v(N) = |L|$ we have $z^b = (1 - z^\ell)|L|/|B|$. Let y be the member of Y in which every buyer's payoff is z^b and every seller's payoff is z^ℓ. We have $y_i = z^b \geq z_i$ and $y_j = z^\ell \geq z_j$, with at least one strict inequality. Further, $y_i + y_j = z^b + z^\ell = (1 - z^\ell)|L|/|B| + z^\ell \leq 1 = v(\{i, j\})$. If we adjust y_i and y_j slightly to make both of the inequalities $y_i \geq z_i$ and $y_j \geq z_j$ strict then $y \succ_{\{i,j\}} z$.

The standard of behavior that this stable set exhibits is "equality among agents of the same type". Note the difference between this set and a set of the type $Y_p = \{x_i = p \text{ for all } i \in L \text{ and } x_j = 1 - p \text{ for } |L| \text{ members of } B\}$ for some p, which can be interpreted as the standard of behavior "the price is p".

280.3 (*Stable sets of three-player games*) The set of imputations is the triangle in Figure 66.1. The core is the heavy line segment at the top of the diagram:

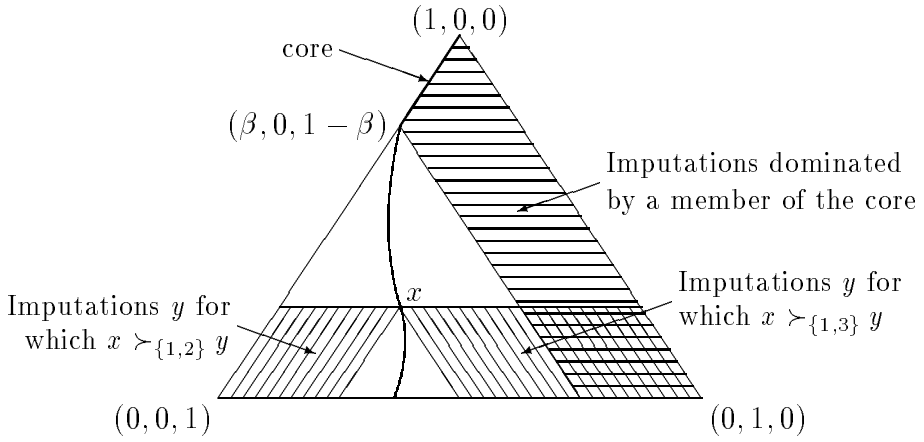

Figure 66.1 The core and a stable set of the three-player game in Exercise 280.3. The triangle is the set of imputations; each corner corresponds to an imputation in which one player obtains a payoff of 1, as labelled. The heavy line at the top of the figure is the core, and the core together with a curved line like the one shown is a stable set. (The curved line extends from $(\beta, 0, 1-\beta)$ to the line $x_1 = 0$ and has the property that all points below any given point on the line lie between the two straight lines through the point parallel to the sloping sides of the triangle.)

the set $\{(\gamma, 0, 1-\gamma): \beta \leq \gamma \leq 1\}$. We know that the core is a subset of every stable set, so that (by internal stability) no imputation that is dominated by a member of the core is a member of any stable set. The set of imputations dominated by a member of the core (via the coalition $\{1,3\}$) is shown in the figure. Now take any of the remaining imputations, say x. The set of imputations that it dominates is the union of the two shaded sets below the horizontal line through it. Thus in order for external stability to be satisfied a stable set must contain every point on some curve joining $(\beta, 0, 1-\beta)$ and the bottom of the triangle. In order for internal stability to be satisfied a stable set can contain only the points on such a line, and the line must have the property that all the points on it below any given point must lie between the two straight lines through the point that are parallel to the sloping sides of the triangle. For example, one stable set consists of the union of the points in the core and the points on the curved line in the figure.

In the core player 2, the buyer with the lower reservation value, obtains nothing. One interpretation of a stable set is that it corresponds to a method of splitting the joint payoff that the buyers can obtain, with the following property: each dollar up to $1 - \beta$ goes entirely to player 3; a nonnegative amount of every subsequent dollar is added to the payoff of both player 2 and player 3.

280.4 (*Dummy's payoff in stable sets*) Let x be an imputation in a stable set and let i be a dummy. Suppose that $x_i > v(\{i\})$ for some $i \in N$. Since $v(N) - v(N \setminus \{i\}) = v(\{i\})$ we have $x(N\setminus\{i\}) < x(N) - v(\{i\}) = v(N) - v(\{i\}) = v(N\setminus\{i\})$, so that there exists an imputation y such that $y \succ_{N\setminus\{i\}} x$. By internal stability we have $y \notin Y$, and thus by external stability there exists an imputation $z \in Y$ and a coalition S with $z \succ_S y$. If $i \notin S$ then we have $z \succ_S x$, contradicting internal stability; if $i \in S$ then $z \succ_{S\setminus\{i\}} x$ since i is a dummy, again contradicting internal stability. Thus $x_i = v(\{i\})$ for all $i \in N$.

280.5 (*Generalized stable sets*) It is straightforward to check that the core is a stable set. It is the only stable set because it must be a subset of any stable set and every imputation not in the core is dominated by an allocation in the core.

283.1 (*Core and bargaining set of market*) Let x be an imputation; without loss of generality assume that $x_1 \leq x_2$ and $x_3 \leq x_4 \leq x_5$. We argue that $x_1 = x_2$ and $x_3 = x_4 = x_5$. Assume not; then either $x_1 < x_2$ or $x_3 < x_5$ and in either case $x_2 + x_5 > 1$. In the arguments below ϵ is a small enough positive number.

If $x_1 + x_3 < 1$ and $x_4 > 1$ then consider the objection $((1 - x_3 - \epsilon, x_3 + \epsilon), \{1,3\})$ of 3 against 2. There is no counterobjection of 2 using either the coalition $\{2,4\}$ (since $x_2+x_4 > 1$) or the coalition $\{2,4,5\}$ (since $x_2+x_4+x_5 > 1$). Adding player 1 to the counterobjecting coalition does not increase its worth. Thus there is no counterobjection to the objection.

If $x_1+x_3 < 1$ and $x_4 \leq 1$ then consider the objection $(y, S) = ((1 - x_3 - 2\epsilon, x_3 + \epsilon, 1 + \epsilon), \{1,3,4\})$ of 3 against 2. If ϵ is small enough there is no counterobjection of 2 using either the coalition $\{2,4\}$ (since $x_2 + y_4 > 1$) or the coalition $\{2,4,5\}$ (since $x_2 + 1 - \epsilon + x_5 > 0$ for ϵ small enough). As before, adding player 1 to the counterobjecting coalition does not increase its worth. Thus there is no counterobjection to the objection.

The remaining case is that in which $x_1 + x_3 \geq 1$. Since $x_2 + x_5 > 1$ we have $x_1+x_3+x_4 < 2$. Consider the objection $((x_1+\epsilon, x_3+\epsilon, 2-x_1-x_3-2\epsilon), \{1,3,4\})$ of 3 against 2. There is no counterobjection of 2 using the coalition $\{2,4\}$ (since $x_2+2-x_1-x_3-2\epsilon > x_2+x_5-2\epsilon$, which, for ϵ small enough, exceeds 1) or the coalition $\{2,4,5\}$ (since $x_2+1-\epsilon+x_5 > 0$. Thus there is no counterobjection to the objection.

We conclude that $x_1 = x_2 = \alpha$ and $x_3 = x_4 = x_5 = \beta$ (say). For any objection of 1 against 2 using the coalition $\{1\} \cup S$ there is a counterobjection of 2 against 1 using the coalition $\{2\} \cup S$. Any objection of 3 against 4 or 5 can be countered similarly. Now consider an objection of 1 against 3. If the

coalition used is $\{1,4\}$ then 3 can counterobject using $\{2,3\}$; if the coalition used is $\{1,4,5\}$ then 3 can counterobject using $\{2,3,4\}$; if the coalition used is $\{1,2,4,5\}$ then 3 can counterobject using $\{2,3,4\}$. By similar arguments any objection of 3 against 1 can be countered.

The core of the game consists of the single imputation $(0,0,1,1,1)$, which is induced by competition between 1 and 2. In any other imputation $(\alpha,\alpha,\beta,\beta,\beta)$ we have $\alpha+\beta<1$, so that a coalition consisting of a seller and a buyer can profitably deviate. According to the reasoning of the players modeled by the bargaining set such a deviation will not occur since whenever one buyer points out that she can get together with a seller and increase her payoff the seller points out that he can get together with another buyer and do so, which convinces the original buyer not to deviate.

289.1 (*Nucleolus of production economy*) Let x be the imputation described in the exercise. We need to show that for every objection (S,y) to x there is a counterobjection T. Let (S,y) be an objection to x. Let $W=N\setminus\{1\}$ (the set of workers).

Suppose that $S\subseteq W$ and $y_i<x_i$ for some $i\in W$. Then $T=\{i\}$ is a counterobjection: $x(T)=x_i>y_i=y(T)$ and $e(T,y)=-y_i>-x_i\geq -|S|x_i=e(S,x)$ (since $x_i=x_j$ for all $i,j\in W$).

Suppose that $S\subseteq W$ and $y_i\geq x_i$ for all $i\in W$. Then $y_1<x_1$; suppose that $y_j>x_j$. We claim that $T=N\setminus\{j\}$ is a counterobjection. We have $x(T)=x(N)-x_j>x(N)-y_j=y(N)-y_j=y(T)$. Further

$$\begin{aligned} e(T,y) &= f(w-1)-(f(w)-y_j)\\ &= y_j-(f(w)-f(w-1))\\ &> x_j-(f(w)-f(w-1))\\ &= -\tfrac{1}{2}(f(w)-f(w-1)) \end{aligned}$$

and $e(S,x)=-\frac{1}{2}|S|(f(w)-f(w-1))\leq -\frac{1}{2}(f(w)-f(w-1))$.

Suppose that $S\ni 1$; let $|S|=s+1$. Since (S,y) is an objection to x we have $y(S)>x(S)$ and $s<w$. We claim that $T=N\setminus S$ is a counterobjection. First note that $y(T)=f(w)-y(S)$ and $x(T)=f(w)-x(S)$, so that $y(T)<x(T)$. We now show that $e(T,y)\geq e(S,x)$, so that T is a counterobjection to (S,y). We have

$$\begin{aligned} e(S,x) &= f(s)-[f(w)-(w-s)\cdot\tfrac{1}{2}(f(w)-f(w-1))]\\ &= f(s)-\frac{2-w+s}{2}f(w)-\frac{w-s}{2}f(w-1) \end{aligned}$$

and

$$\begin{aligned} e(T,y) &= -y(T) \\ &> -x(T) \\ &= -(w-s)\cdot\tfrac{1}{2}(f(w)-f(w-1)) \\ &\geq f(s) - \frac{2-w+s}{2}f(w) - \frac{w-s}{2}f(w-1), \end{aligned}$$

since by the concavity of f we have $f(w)-f(s) \geq (w-s)(f(w)-f(w-1))$.

289.2 (*Nucleolus of weighted majority games*) We do not have any direct solution to this exercise. (The result is taken from Peleg (1968), who provides a proof based on the standard definition of the nucleolus.)

294.2 (*Necessity of axioms for Shapley value*)

a. The arguments for DUM and ADD are the same as those for the Shapley value. The value ψ does not satisfy SYM: let $N = \{1,2\}$ and consider the game v defined by $v(\{1,2\}) = 1$ and $v(\{1\}) = v(\{2\}) = 0$. Players 1 and 2 are interchangeable but $\psi_1(v) = 0$ and $\psi_2(v) = 1$.

b. The value ψ clearly satisfies SYM and ADD. It does not satisfy DUM: let $N = \{1,2\}$ and consider the game v defined by $v(\{1,2\}) = v(\{1\}) = 1$ and $v(\{2\}) = 0$. Player 2 is a dummy but $\psi_2(v) = \frac{1}{2} \neq v(\{2\})$.

c. The value ψ clearly satisfies SYM and DUM. The following example shows that it does not satisfy ADD. Let $N = \{1,2\}$ and define v by $v(\{1\}) = 0$ and $v(\{2\}) = v(\{1,2\}) = 1$ and w by $w(\{1\}) = w(\{2\}) = 0$ and $w(\{1,2\}) = 1$. Then player 1 is a dummy in v, so that $\psi_1(v) = 0$, while $\psi_1(w) = \frac{1}{2}$; we find that $\psi_1(v+w) = 1 > \psi_1(v) + \psi_1(w)$.

295.1 (*Example of core and Shapley value*) The core is $\{(1,1,1,0)\}$ since for any $1 \leq i < j \leq 3$ we need $x_i + x_j \geq v(\{i,j\}) = 2$ in order for x to be in the core.

The Shapley value gives player 4 a payoff of $\frac{1}{4}$ since his marginal contribution is positive only in orderings in which he is last, and it is 1 in such an ordering. The other players are symmetric, so that the Shapley value of the game is $(\frac{11}{12}, \frac{11}{12}, \frac{11}{12}, \frac{1}{4})$.

Player 4 obtains a payoff of 0 in the core, despite the fact that his presence makes a difference to the amount of payoff that the other players can obtain. The reason is that the core is sensitive to the demands of the two-player coalitions among players 1, 2, and 3, each of which can obtain a payoff

of 2 and player 4 needs at least two of these players to obtain a positive payoff. The Shapley value, on the other hand, takes into account the "marginal contribution" of each player to each possible coalition.

295.2 (*Shapley value of production economy*) The Shapley value gives player 1 (the capitalist) a payoff of $\sum_{i=1}^{w} f(i)/(w+1)$ since in any ordering of the players in which she follows i workers her marginal contribution is $f(i)$ and the probability of her following i workers is $1/(w+1)$. The workers are symmetric, so the Shapley value gives each of them a payoff of $(f(w) - \sum_{i=1}^{w} f(i)/(w+1))/w$.

295.4 (*Shapley value of a model of a parliament*)

a. Let the two large parties be players 1 and 2. If n is large then each of the following sets of orderings has probability close to $\frac{1}{4}$. A: Players 1 and 2 are both in the first half of the ordering; B: Players 1 and 2 are both in the second half of the ordering; C: Player 1 is in the first half of the ordering and player 2 in the second half; D: Player 1 is in the second half of the ordering and player 2 is in the first half. The marginal contribution of player 1 is 0 except in orderings in A in which she comes after player 2 and in orderings in B in which she comes before player 2, in which cases it is 1. Thus her expected contribution is $\frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{4}$.

b. The total share of the small parties is $\frac{1}{2}$ if they are independent; if they unite then the game is symmetric and they obtain only $\frac{1}{3}$.

295.5 (*Shapley value of convex game*) This follows from the result in Exercise 260.4, the definition of the Shapley value, and the convexity of the core.

296.1 (*Coalitional bargaining*)[1]

First we show that the strategy profile in which each player $i \in S$ proposes $x^{i,S}$ whenever the set of active players is S and each player j accepts a proposal y of player i when the set of active players is S if and only if $y_j \geq x_j^{S,i}$ is a subgame perfect equilibrium. It is immediate that this acceptance rule is optimal. To show that player j's proposals are optimal note that by proposing $x^{S,j}$ he obtains $x_j^{S,j}$; any proposal that gives him a higher payoff is rejected, so that he obtains $\rho \overline{x}_j^S + (1-\rho)v(\{j\})$. Thus to complete the argument we need

[1] *Correction to first printing*: After "active players" on line 5 add ", initially N,".

to show that $\rho\overline{x}_j^S + (1-\rho)v(\{j\}) \leq x_j^{S,j}$, or

$$\rho\overline{x}_j^S + (1-\rho)v(\{j\}) \leq v(S) - \rho\sum_{k\in S\setminus\{j\}}\overline{x}_k^S - (1-\rho)\sum_{k\in S\setminus\{j\}}\overline{x}_k^{S\setminus\{j\}}$$

or

$$\rho\sum_{k\in S}\overline{x}_k^S + (1-\rho)v(\{j\}) \leq v(S) - (1-\rho)\sum_{k\in S\setminus\{j\}}\overline{x}_k^{S\setminus\{j\}}.$$

Now, $\sum_{k\in S}\overline{x}_k^S = v(S)$ and $\sum_{k\in S\setminus\{j\}}\overline{x}_k^{S\setminus\{j\}} = v(S\setminus\{j\})$, so that the inequality follows from the assumption that $v(S\cup\{i\}) \geq v(S) + v(\{i\})$ for every coalition S and player $i \in N\setminus S$.

To show that there is a subgame perfect equilibrium for which $\overline{x}^S = \varphi(S,v)$ for each $S\in\mathcal{C}$, let $x_j^{S,i} = \rho\varphi_j(S,v) + (1-\rho)\varphi_j(S\setminus\{i\},v)$ for each coalition S, $i\in S$, and $j\in S\setminus\{i\}$ and $x_i^{S,i} = v(S) - \sum_{j\in S\setminus\{i\}}x_j^{S,i}$. We have $\sum_{j\in S\setminus\{i\}}x_j^{S,i} = \rho(v(S)-\varphi_i(S,v)) + (1-\rho)v(S\setminus\{i\})$, so that $x_i^{S,i} = (1-\rho)(v(S)-v(S\setminus\{i\})) + \rho\varphi_i(S,v)$. Further, using the fact that the Shapley value satisfies the balanced contributions property we have $x_j^{S,i} = \varphi_j(S,v) - (1-\rho)(\varphi_i(S,v) - \varphi_i(S\setminus\{j\},v))$ for $j\in S\setminus\{i\}$. Thus

$$\begin{aligned}\sum_{i\in S}x_j^{S,i} &= (|S|-1)\varphi_j(S,v) - (1-\rho)(v(S)-\varphi_j(S,v)) + \\ &\qquad\qquad (1-\rho)v(S\setminus\{j\}) + x_j^{S,j} \\ &= |S|\varphi_j(S,v),\end{aligned}$$

so that $\overline{x}^S = \varphi(S,v) = \sum_{i\in S}x^{S,i}/|S|$ as required.

15 The Nash Bargaining Solution

309.1 (*Standard Nash axiomatization*) See, for example, Osborne and Rubinstein (1990, pp. 13–14).

309.2 (*Efficiency vs. individual rationality*) Fix $\alpha \in (0,1)$ and consider the solution F defined by $F(X, D, \succsim_1, \succsim_2) \sim_i \alpha \cdot N(X, D, \succsim_1, \succsim_2)$ for $i = 1, 2$, where N is the Nash solution.

Strict individual rationality: This follows from the fact that the Nash solution is strictly individually rational.

SYM: Suppose that $\langle X, D, \succsim_1, \succsim_2 \rangle$ is symmetric, with symmetry function ϕ. Since $F(X, D, \succsim_1, \succsim_2) \sim_i \alpha \cdot N(X, D, \succsim_1, \succsim_2)$ we have

$$\phi(F(X, D, \succsim_1, \succsim_2)) \sim_j \phi(\alpha \cdot N(X, D, \succsim_1, \succsim_2)) \text{ for } j \neq i.$$

But

$$\phi(\alpha \cdot N(X, D, \succsim_1, \succsim_2)) = \alpha \cdot \phi(N(X, D, \succsim_1, \succsim_2)) = \alpha \cdot N(X, D, \succsim_1, \succsim_2).$$

Thus $\phi(F(X, D, \succsim_1, \succsim_2)) \sim_i F(X, D, \succsim_1, \succsim_2)$ for $i = 1, 2$. Finally, from the non-redundancy assumption we have $\phi(F(X, D, \succsim_1, \succsim_2)) = F(X, D, \succsim_1, \succsim_2)$.

IIA: Since $N(X, D, \succsim_1', \succsim_2) = N(X, D, \succsim_1, \succsim_2)$ for preference relations that satisfy IIA we have $F(X, D, \succsim_1', \succsim_2) \sim_i F(X, D, \succsim_1, \succsim_2)$. Thus from the non-redundancy assumption F satisfies IIA.

What accounts for the difference between Roth's result and the one here is that Roth's argument uses a comparison between two bargaining problems with different sets of agreements, while here the set of agreements is fixed.

310.1 (*Asymmetric Nash solution*)

Well-definedness: Suppose that u_i and v_i both represent player i's preferences, for $i = 1, 2$. Then $u_i = \gamma_i v_i + \beta_i$ for some $\gamma_i > 0$ for $i = 1, 2$, so

that $(v_1(x) - v_1(D))^\alpha (v_2(x) - v_2(D))^{1-\alpha} = \gamma_1^\alpha \gamma_2^\alpha (u_1(x))^\alpha (u_2(x))^{1-lpha}$. Thus the asymmetric Nash solution is well-defined.

PAR: This follows from the definition of the solution as the maximizer of an increasing function on X.

IIA: Let F be an asymmetric Nash solution. Suppose that $\succsim_1'$ satisfies the hypotheses of IIA and let u_1, u_2, and v_1 represent the preferences $\succsim_1$, $\succsim_2$, and $\succsim_1'$ respectively with $u_1(D) = u_2(D) = v_1(D) = 0$. We claim that $F(X, D, \succsim_1, \succsim_2) = F(X, D, \succsim_1', \succsim_2)$. Suppose, to the contrary, that $x^* = F(X, D, \succsim_1, \succsim_2)$ is not the asymmetric Nash solution of $\langle X, D, \succsim_1, \succsim_2 \rangle$. Then there exists $x \in X$ such that $(v_1(x))^\alpha (u_2(x))^{1-\alpha} > (v_1(x^*))^\alpha (u_2(x^*))^{1-\alpha}$, or $(u_2(x)/u_2(x^*))^{1-\alpha} > (v_1(x^*)/v_1(x))^\alpha$. Now, since x^* is the asymmetric Nash solution of $\langle X, D, \succsim_1, \succsim_2 \rangle$ we have $(u_1(x))^\alpha (u_2(x))^{1-\alpha} \geq (u_1(x^*))^\alpha (u_2(x^*))^{1-\alpha}$, or $(u_1(x^*)/u_1(x))^\alpha \geq (u_2(x)/u_2(x^*))^{1-\alpha}$. It follows that $u_1(x^*)/u_1(x) > v_1(x^*)/v_1(x)$. Thus if $x \succsim_1 x^*$ and $p \cdot x \sim_1 x^*$ then $p = u_1(x^*)/u_1(x) > v_1(x^*)/v_1(x)$, so that $p \cdot x \succ_1' x^*$, violating the hypotheses about $\succsim_1'$ in IIA.

Differs from Nash solution: Suppose that the preferences are such that $\{(u_1(x), u_2(x)) : x \in X\}$ is the convex hull of $(0,0)$, $(1,0)$, and $(0,1)$. Then the Nash solution yields the pair of utilities $(\frac{1}{2}, \frac{1}{2})$ while an asymmetric Nash solution with parameter α yields the utilities $(\alpha, 1-\alpha)$.

310.2 (*Kalai–Smorodinsky solution*)

Well-definedness: This is immediate from the definition.

PAR: This is immediate from the definition.

SYM: Let $\langle X, D, \succsim_1, \succsim_2 \rangle$ be a symmetric bargaining problem with symmetry function ϕ. Let x^* be the Kalai–Smorodinsky solution of $\langle X, D, \succsim_1, \succsim_2 \rangle$. We need to show that $\phi(x^*) = x^*$. First we argue that $\phi(x^*)$ is Pareto efficient. Suppose to the contrary that there exists $x \in X$ such that $x \succ_i \phi(x^*)$ for $i = 1, 2$. Then from the definition of a symmetric bargaining problem we have $\phi(x) \succ_j \phi(\phi(x^*)) = x^*$ for $j = 1, 2$, contradicting the Pareto efficiency of x^*. We now claim that $u_1(\phi(x^*))/u_2(\phi(x^*)) = u_1(B_1)/u_2(B_2)$. Since x^* is the Kalai–Smorodinsky solution of $\langle X, D, \succsim_1, \succsim_2 \rangle$ we have $u_1(x^*)/u_1(B_1) = u_2(x^*)/u_2(B_2) = p \leq 1$, so that $x^* \sim_1 p \cdot B_1$ and $x^* \sim_2 p \cdot B_2$. Therefore by the symmetry of the bargaining problem we have $\phi(x^*) \sim_2 p \cdot \phi(B_1) = p \cdot B_2$ and $\phi(x^*) \sim_1 p \cdot \phi(B_2) = p \cdot B_1$, so that $u_1(\phi(x^*))/u_2(\phi(x^*)) = u_1(B_1)/u_2(B_2)$ and hence $\phi(x^*)$ is a Kalai–Smorodinsky solution of $\langle X, D, \succsim_1, \succsim_2 \rangle$. Thus $\phi(x^*) = x^*$.

Differs from Nash solution: Let $d = (u_1(D), u_2(D))$ and suppose that $S = \{(u_1(x), u_2(x)) : x \in X\}$ is the convex hull of $(0,0)$, $(1,0)$, $(\frac{1}{2}, \frac{1}{2})$, and $(0, \frac{1}{2})$.

The Kalai–Smorodinsky solution is the x^* for which $(u_1(x^*), u_2(x^*)) = (\frac{2}{3}, \frac{1}{3})$ while the Nash solution is the x' for which $(u_1(x^*), u_2(x^*)) = (\frac{1}{2}, \frac{1}{2})$.

312.2 (*Exact implementation of Nash solution*) **Note**: In the first and second printings of the book it is suggested that the proof follow three steps.[1] However, a shorter proof, not following the steps, can be given as follows.

First note that if player 1 chooses x^* at the first stage then player 2 can do no better than choose $(x^*, 1)$ at the second stage. This follows since the outcome is either $p \cdot x$ or $p^2 \cdot x^*$ (where (x, p) is the choice of player 2 at the second stage), and if $p \cdot x \succ_2 x^*$ then from the definition of the Nash solution (301.2) we have $p \cdot x^* \succ_1 x$, so that the outcome is $p^2 \cdot x^*$. Thus all subgame perfect equilibrium outcomes are at least as good for player 1 as x^*.

Now, let y be the choice of player 1 in the first stage. By choosing (x, p) for which $x \succ_1 p \cdot y$ in the second stage player 2 can obtain the outcome $p \cdot x$. Letting u_i for $i = 1, 2$ be a von Neumann–Morgenstern utility function that represents $\succsim_i$ and satisfies $u_i(D) = 0$, this means that for any $p < u_1(x)/u_1(y)$ player 2 can achieve the payoff $pu_2(x)$. Thus in an equilibrium player 2's payoff is equal to $\max_{x,p} pu_2(x)$ subject to $p \leq \min\{u_1(x)/u_1(y), 1\}$. If $u_1(y) > u_1(x^*)$ then the solution of this problem is $(x, p) = (x^*, u_1(x^*)/u_1(y))$, in which case player 1's payoff is less than $u_1(x^*)$. If $u_1(y) < u_1(x^*)$ then the solution of the problem is $(x, p) = (x, 1)$ where $x \sim_1 y$; thus player 1's payoff is $u_1(y)$. Since player 1's payoff in equilibrium is $u_1(x^*)$, neither case is thus an equilibrium. Finally, if $u_1(y) = u_1(x^*)$ but $y \neq x^*$ then player 2 chooses $(x^*, 1)$ and the outcome is x^*. Thus in any subgame perfect equilibrium the outcome is x^*.

(Note that in addition to the equilibrium in which player 1 chooses x^* and player 2 chooses $(x^*, 1)$, for any y with $y \prec_1 x^*$ there is an equilibrium in which player 1 chooses x^* and player 2 chooses $(y, 1)$.)

[1] These steps require slight modifications: for example, if in Step 1 y is efficient then we can conclude only that either $p < 1$ and $p \cdot y \sim_1 x$, or $p = 1$ and $p \cdot y \succsim_1 x$.